

Rules of genome access

Statements by presidents of countries and societies highlight the concern that human genome data be publicly accessible, and quickly. The devil may lie in the emerging terms of public access to privately owned databases.

As many commentators have noted, there is less than meets the eye in the content of recent statements by President Bill Clinton and Prime Minister Tony Blair that the raw data from human genome projects should be released rapidly into the public domain. But they affect the climate in which decisions over patents and investment are taken. Now the presidents of the National Academy of Sciences and of the Royal Society have issued a joint statement in this week's issue (see page 325) urging that the human genome sequence "must be freely available to all humankind", and that patents should only be awarded to applications in which significant utility is demonstrated.

The principal, if unmentioned, target of these concerns is Celera Genomics, the private rival of the publicly funded human genome project. The existence of the public project is a guarantee against a potentially catastrophic monopoly situation. Celera's financial interests lie in its combination of intellectual property derived from genes and their variability, and from charging subscribers for access to software that accurately pulls out genes and makes good predictions of function by integrating information from a range of sources, including sequence homologies in different databases and relevant literature. The value of such services is reflected in the \$283 million acquisition by Celera of the software company Paracel.

But what are Celera's terms of access? In statements to *Nature*, the company has promised that academic researchers will have free access at the time of publication — some months after subscribers get unconditional access — to the primary sequence data from its

website, or on DVD, without any restriction, and can publish and seek patents on discoveries without any 'reach through' rights (see page 324). Anyone, except competitors, will be able, for example, to use the data to develop DNA chips, it claims.

The sticking point is annotation — the identification of gene sequences and assigning probable function. Traditionally, researchers have annotated sequences and made the results available freely. But it appears that Celera would not allow researchers who are not subscribers to take the whole genome, annotate it and distribute it in competition. This restriction appears unreasonable, given that the challenge of making good predictions of function is in the software and not in the sequence itself.

Celera's ambitions highlight the potentially changing rules of the game of databases. There is every possibility that large data sets will increasingly be held by privately funded ventures.

Nature's policy is that human sequence data should be deposited in a reliable, publicly available, unrestricted and free database. GenBank and its equivalents in Europe and Japan satisfy those conditions. But there is no insistence in principle that deposition be in databases that are publicly funded — the conditions of access and the community's confidence in the long-term sustainability and accessibility of the database are what count. The value of databases such as GenBank lies partly in their role as an honest broker of information for the community at large, and any private owner of a database would do well to have earned a strong reputation in that respect too. ■

Eyes, ears and science

Two artists have succeeded in involving two or more senses to produce something special.

Most of us are likely to approach new art in ignorance of current aesthetic theories and trends and of the artist's previous output, in the hope that it will speak to us or at least grip our attention. Statistically, this hope is likely to be forlorn — innovative artists with an ability immediately to communicate beyond the cognoscenti are a minority in any generation. It's surprising and rewarding, therefore, to encounter artists working in unconventional media who make one sit up and take notice.

Two such artists are Carsten Nicolai, based in East Berlin, and the Danish artist Ann Lislegaard, who have recently exhibited at the Museum of Modern Art in Oxford, UK, and have also exhibited widely in the United States and elsewhere in Europe. The Oxford exhibition focused on works combining sound and vision, and these two artists made an impact for reasons associated with science.

Lislegaard's *In another room* consists of a large white wall towards which is pointed a 500-watt white halogen light. A recording of two people — a man moving around in a room, a woman describing his actions — plays softly in the background. The light illuminates the wall with a rapidly varying intensity proportional to that of the sound. This leads to an effect that challenges one's sense of orientation. Whatever the neural mechanisms at work, contemplating the wall for sever-

al minutes produces sensations that are unexpected and memorable.

The interest in Nicolai's *Atem* is less perceptual — it is directly related to known physics — but is still striking and even beautiful. A raised wooden floor has eight large speakers embedded in it, all sounding quietly at slightly different low frequencies — a few hertz — generating both an audible and vibrational effect on the onlooker. The frequencies are shifted slowly, and the envelope of sound amounts to a gradually changing pattern of pulsation, only just discernible as such. But Nicolai places two large glass flasks on the floor, illuminated at an angle by a spotlight. The floor's motion is picked up by water in the flasks, which refract the light onto the floor in a pattern affected by the water waves set up by the vibrations. As the audible frequencies shift, the patterns of light exhibit slowly changing networks of cusps of illumination similar to, but much more fascinating than, those sometimes seen on a swimming-pool floor in bright sunlight. Again, the cumulative effect, drawing on three of the senses, is more than the sum of the parts, each of which is striking in itself.

These works deserve wide exposure, and could in principle be easily replicated in galleries and science museums. But are they works of art or sophisticated, science-based entertainment? That question is left as an exercise for the reader, best answered only after experiencing them. ■





Space projects
US rules 'hinder international efforts at cooperation'
p321



Automakers out
General Motors leaves anti-Kyoto lobby group
p322



Synchrotron site
Daresbury loses out to the south in UK location controversy
p323



Genome data
Celera reveals the terms of access to its data bank
p324

Researchers caught in dispute over transgenic mice patents

San Diego

Neuroscientists who use mouse models are closely watching a patent infringement battle between a prominent biotechnology firm and a major US non-profit institution.

Elan Pharmaceuticals, of Ireland, is suing the Minnesota-based Mayo Foundation for Medical Education and Research over mice used to study Alzheimer's disease. The company alleges that Mayo has infringed two patents it holds on transgenic mice with a mutation that causes Alzheimer's.

Elan uses its closely guarded mice for work on potential drugs against the neurodegenerative disease. The company alleges that Mayo is infringing its patents by making, using, and selling mice that overexpress an Alzheimer amyloid precursor protein, APP695, from a human genetic mutation. Mayo says it has licences from patent holders for both the mutation and the mice, which it distributes to researchers worldwide. A jury is due to decide the issue in October.

According to documents filed with the US District Court in San Francisco, Mayo is making the counterclaim that the Elan patents were obtained fraudulently from the US Patent and Trademark Office and that they are unenforceable and therefore invalid. Mayo attorney Karen Boyd would not specify the grounds for such allegations.

The battle is spreading unease among US neuroscientists because some are being subpoenaed — along with their laboratory notebooks — to testify in depositions for the civil lawsuit quietly filed by Elan last April.

"It is outrageous," says Karen Duff, a New York University neuroscientist who has been subpoenaed by Elan. After receiving her doctorate at Imperial College London, Duff once worked at Mayo's research facility in Jacksonville, Florida, and has used the Mayo mouse model to create another transgenic mouse that is considered extremely important for research.

The dispute is seen by neuroscientists as complicating an already difficult research arena, in which the number of animal models for unravelling the biological secrets of the disease is already limited.

"All of this does no one any good," says

John Trojanowski, a physician with a doctorate in neuroscience who directs a National Institute of Aging (NIA) Alzheimer Center at the University of Pennsylvania. "The Alzheimer research field could move considerably faster" without the lawsuit, he argues.



Younkin: licensing agreement necessary.

In December 1998, as the Elan-Mayo dispute was heating up, NIA officials held a private two-day workshop of neuroscientists to address problems associated with Alzheimer's research using transgenic rodents.

Although Mayo and Elan subsequently held discussions about resolving their dispute out of court, Elan filed its lawsuit a few months later.

Elan attorney Jean Duvall says that no one at the firm's Dublin headquarters will comment except to say "it is Elan's policy to enforce its intellectual property rights".

The Elan-Mayo lawsuit is so sensitive that Marcelle Morrison-Bogorad, NIA's

associate director for neuroscience and neuropsychology, cancelled a scheduled interview on the subject last week — shortly after the joint statement by President Bill Clinton and British Prime Minister Tony Blair on the need for open access to genetic data (see page 324).

"It is critically important that research tools are available to scientists," Morrison-Bogorad says. "We are looking into the current controversy over research resources for Alzheimer's disease and evaluating its impact on the field, on initiatives in the private sector and on the overall progress of research."

The Elan-Mayo legal dispute involves a number of patent, licensing and research agreements with both proprietary and academic institutions. The two patents Elan alleges are being infringed are owned jointly by the firm and Eli Lilly & Co.

Elan secured rights to the patents in 1998, after acquiring Athena Neurosciences, a firm that had licensed research findings from academic institutions and companies.

The two Elan patents are for "transgenic... rodents" that harbour an allele for the 'Swedish mutation' — named after a large

Iridium crash relieves astronomers

Munich

Radioastronomers may not be gloating publicly over the final plunge of the global mobile telephone company Iridium into bankruptcy last week. But news that the company's 66 satellites must be brought out of orbit and burnt up in the atmosphere as part of the bankruptcy agreement has ignited a private outbreak of *schadenfreude*.

Radioastronomers have been fighting for years against the satellite system developed by Iridium, which was allocated a waveband right beside the 1612 MHz band reserved for astronomical observations. Despite assurances, Iridium refused to cap the use of its satellites at a level that would avoid 'overspill' into the radioastronomy band, drowning out weak radio signals from space.

Last year, European radioastronomers reached a compromise with Iridium under which the company was required to guarantee some hours of 'quiet time' over radio observatories (*Nature* 399, 513; 1999). But since their launch in November 1998, the 66 brightly reflecting satellites have also interfered with optical observations from ground-based telescopes.

"We have learnt from the affair that promises made at one time are not always kept," says Willem Baan, director of the Westerbork Observatory in northern Holland and one of the fiercest opponents of Iridium. Baan is now working to ensure that, in future, frequencies reserved for science are protected from unexpected commercial encroachment.

Allison Abbott

► Scandinavian family in which it was identified — that is believed to contribute to Alzheimer's disease by producing a build-up of amyloid plaque in the brain.

Researchers say that Elan transgenic mice are closely held by the firm for its drug development efforts. One researcher alleges that when Elan mice are provided to academic scientists, they are usually neutered females.

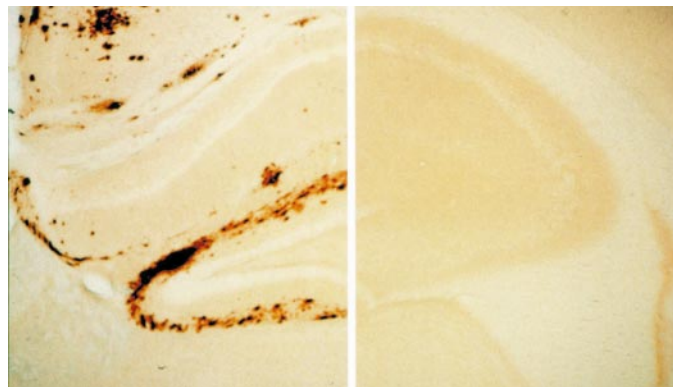
Mayo's transgenic mice, provided to more than 50 academic research groups and a dozen pharmaceutical firms, are based on work by a group headed by Karen Hsiao at the University of Minnesota.

Hsiao's group created a mouse that would express amyloid in the brain (see *Science* 274, 99–103; 1996). After publication, Mayo licensed the discovery from the University of Minnesota, arranged for contract breeding of the mice and began distributing them. Mayo also licensed the sequence of the Swedish mutation from a Kansas firm.

Mayo's mice are considered so valuable that there are reports of breeding trios of males being sold to companies for \$850,000. Mayo officials decline to discuss this.

When providing its mice to academic institutions, Mayo requires the signing of a material transfer agreement that gives the non-profit organization an option to buy the rights to any commercial discovery that may come from research with the mice.

This agreement — like those used by biotech firms — is considered by some to be an unusually bold move for a non-profit organization. Even some of Mayo's own



Mice work: the brain of Elan's one-year-old mouse (far left) shows characteristic plaques of Alzheimer's disease, absent from that of the similar animal (left), which was dosed with Elan's experimental Alzheimer's vaccine.

researchers have difficulty with this practice.

"My view is there should be no reach-through agreements between non-profit institutes," says pharmacology professor John Hardy at Mayo's Jacksonville facility, whose laboratories at Imperial College London, South Florida and Mayo have produced leading discoveries and researchers.

But Steven Younkin, a physician neuroscientist and former director of research at Mayo's Jacksonville facility, defends the agreements as being necessary to cover the organization's enormous cost for the long-term, mouse-producing project. "With a non-profit institution such as Mayo, any money realized from licensing agreements goes back into research," says Younkin.

If Mayo was primarily interested in making money, he adds, it would have entered an exclusive licensing agreement for the transgenic mice with a single pharmaceutical

firm. It deliberately decided not to do this, in order to make the mice available for academic research.

As neuroscientists debate these issues, subpoenas have been spreading through the neuroscience community. Earlier this month, Elan failed in a bid to require Hsiao to produce her laboratory notebooks for a deposition this week.

A number of researchers, including microbiologist David Borchelt, an associate professor of pathology at Johns Hopkins University in Maryland, are debating their options as they face subpoenas. Unable to secure the desired transgenic mice, Borchelt made his own mice.

Elan has now subpoenaed him and his lab notebooks, which he fears may be studied closely by the firm's scientists. Borchelt says he would "go to jail" rather than provide his notebooks. **Rex Dalton**

Computer glitch unleashes prize nomination debate

Boston

The normally tranquil world of statistical physics and mechanics was rocked last week by a computer glitch that led to a flurry of exchanges over nominations for the field's most prestigious prize, the Boltzmann Medal. The prize is awarded by the International Union of Pure and Applied Physics (IUPAP) once every three years.

The glitch triggered a debate between those who regarded the proliferating e-mail messages as a nuisance that threatened to overwhelm computer networks, and others who welcomed an unaccustomed openness.

The source of the problem was the 3,500-strong mailing list kept by the Center for Mathematical Sciences Research at Rutgers University, which hosts two conferences on statistical mechanics each year.

Kurt Binder of the University of Mainz in Germany, who chairs the IUPAP committee that issues the Boltzmann Medal, asked Joel Lebowitz — a member of the nominating committee — to send an announcement soliciting nominations. (The next prize will

be awarded in Mexico, in summer 2001.)

Recipients were meant to reply to him alone, not sending their comments to each other.

Juerg Froehlich of Eidgenössische Technische Hochschule Zurich nominated John Cardy of Oxford University, also citing the work of Sasha Zamolodchikov of Rutgers. "But something went wrong," says Lebowitz. "The computer went haywire" when Froehlich hit the 'reply' button, and his suggestions went to all 3,500 on the list.

Others soon joined in the nomination frenzy. Roger Bidaux of the Centre d'Etudes de Saclay in France advanced the name of Lawrence Schulman, a physicist at Clarkson University in Potsdam, New York. "I believe that our community can afford a little bit of tolerance and democracy, and allow me to express my opinion," wrote Bidaux.

University of Rochester physicist Yonathan Shapir shared this sentiment, saying he'd like to see the process conducted in a more open, democratic fashion.

"There is a lot of secrecy in academia," says Shapir. "I don't like the way this business

is handled, with just a few people making decisions for the rest of us." The Internet, he adds, "could allow us to hold large 'town meetings,' where many people voice their opinions and, in the end, a vote is taken".

Several respondents, though, argued that nominations should not be aired in public.

Meanwhile, Haym Benaroya of Rutgers was anxious to avert an e-mail meltdown, and urged people to stop sending messages to the Rutgers list. "At one point, I was getting 10 to 15 e-mails every 10 minutes," he says. "With everybody talking to everybody else, things were snowballing. The whole system could have been shut down."

The snafu was corrected within a day, and an apology issued. Future nominations were requested by post, rather than e-mail.

The apology put an official end to the episode. But Shapir hopes that an important lesson might still be learned. As a result of the computer glitch, he says, "for a brief period of time, we got a glimpse of how things could be done. We don't have to cling to the secretive ways of the past." **Steve Nadis**

Japan's cloning ban will still allow stem cell experiments

Tokyo

Japan's parliament, the Diet, is expected to approve a new law later this month prohibiting the reimplantation into the uterus of human clones, chimaeras, or 'hybrids' — but allowing experimentation with cloned embryos, subject to strict guidelines.

The text of the law has been drawn up by the Science and Technology Agency (STA) after consultation with an expert panel. In contrast to the regulations covering federally funded research in the United States, some forms of nuclear transfer will be allowed.

The relatively limited scope of the legal restrictions has surprised members of the committees drafting the bill as well as some observers in Tokyo. In particular, it appears that the reimplantation of clones that have been created using cells from human embryos — defined as an 'unborn' human being — will not be banned, but merely subject to guidelines and a review of experimental protocols.

STA officials say that the new bill does justice to both science and ethics by prohibiting human cloning without preventing potential opportunities that may emerge from stem cell and human embryo research. Guidelines covering the latter are to be released later this month by the Council of Science and Technology, an advisory body close to STA.

But critics argue that the new law could open a 'back door' for human cloning research. And members of the expert panel say that the distinction made by STA between the nuclear transfer from cells of 'born' and 'unborn' human beings "had never been discussed at the panel".

Sources in Tokyo say that interventions by politicians and the medical community, worried that tight regulations could limit future economic opportunities for Japanese companies in such areas as stem cell research, have greatly influenced the final shape of the bill.

Motoya Katsuki, a professor in embryology at the Institute of Medical Research at the University of Tokyo and a member of the cloning and human embryo research expert panels, accepts that using guidelines rather than legal constraints in areas of science that are changing fast will make regulation more flexible. But he argues that the resulting regulatory framework is clumsy and may lack credibility in the eyes of the public.

Robert Triendl

Space researchers protest at 'disruptive' export controls

Washington

US space scientists are becoming increasingly alarmed about strict new government export control policies introduced in response to Congressional concern over the sales of missile technology to China.

A Hungarian scientist, for example, who built an instrument for a satellite being developed by Stanford University, was denied access to inspect his own instrument after it entered the United States.

"It flies in the face of reason," says Bradford Parkinson of Stanford, chairman of the advisory council to the space agency NASA.

The new regulations place an "impossible burden on most universities" working on spacecraft missions, according to Claude Canizares of the Massachusetts Institute of Technology (MIT), who chairs the National Academy of Sciences' Space Studies Board.

In 1998 Congress transferred authority over satellite technology exports from the Commerce Department to the State Department, which administers the International Traffic in Arms Regulations (ITAR). The shift has effectively placed researchers working on scientific spacecraft in the same category as munitions exporters.

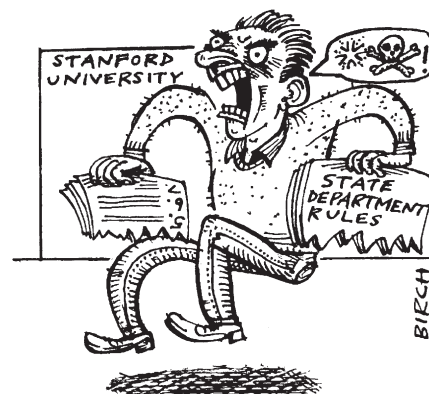
Although ITAR exempts scientific research, NASA policy directives issued last month do not. This has caused widespread concern among space scientists who say they are now unsure whether foreign-born graduate students are allowed to work on satellite hardware and data without first obtaining an export licence.

Canizares aired his concerns at a meeting last week of the NASA Advisory Council. His move prompted other academic scientists on the panel to share horror stories about projects mired in red tape and legal uncertainty.

Steven Squyres of Cornell University, who plays a leading role in the US–French Mars exploration programme, said that ITAR issues are particularly irksome in NASA projects that are trying to operate in a 'better, faster, cheaper' mode. "The more international we get, the more of a problem it's going to be," he told the advisory council.

Canizares and Eugene Skolnikoff of MIT, who chairs an academy committee on international space programmes, wrote last month to National Research Council chairman Bruce Alberts, suggesting that the academy conduct a study of the ITAR problem.

Other groups, including the Association of American Universities and the Council on Governmental Relations (COGR), a Washington-based association of research universities, have also taken up the cause, and have been working with several universities —



IN INTERNATIONAL SPACE PROJECTS, NOBODY CAN HEAR YOU SCREAM...

such as Stanford, MIT, the University of California, and the University of Maryland — on a plan of action. The NASA Advisory Council also formed a subcommittee last week to look into the issue, led by Canizares.

Tony DeCrappeo of COGR says that the problems being experienced by NASA-funded researchers will probably extend to other agencies, including the Department of Energy and the Department of Defense. But he believes that an answer lies within reach if existing laws are interpreted properly. "I think the exception [for researchers] as written in the ITAR is probably okay," he says, meaning that no new legislation would have to be written to exempt scientists.

But all agree that a clarification of the policy has to come from the top. A similar confusion over regulations governing foreign access to very-high-speed integrated-circuit technology in the 1980s was settled after years of debate by a national security directive from President Ronald Reagan.

NASA Administrator Daniel Goldin, who heard the scientists' complaints at the meeting, urged them not to "get preachy," and to respect legitimate national security concerns — particularly when the current political climate in Washington favours a 'get tough' attitude toward the leaking of high-tech secrets. Congress recently passed a law preventing NASA from sending more money to Russia for the international space station project unless questions about transfer of missile technology to Iran are resolved.

"It is a very difficult environment now," said Goldin, who is nevertheless eager to reassure the university community, which he envisages will play a larger role in NASA-sponsored research as the agency's budget increases. "The transformation of NASA cannot occur without a stronger relationship with universities," he told the advisory council.

Tony Reichhardt

Car maker joins exodus from anti-Kyoto coalition

Washington

The US company General Motors (GM) last week became the third major automobile manufacturer to withdraw from a lobbying group that has been leading opposition to the 1997 Kyoto Protocol, an accord designed to reduce global warming.

The move brings to five the number of large corporations that have recently resigned from the Global Climate Coalition (GCC), an 11-year-old industry coalition that vigorously opposes the Kyoto Protocol.

GM told the group last week that it was not renewing its membership. Its withdrawal follows that of Ford in December and DaimlerChrysler in January. The oil company Texaco and The Southern Company, a large US power company, have also recently resigned (see *Nature* 404, 118; 2000.)

In response, the GCC announced that it would no longer allow individual companies as members, but only trade associations, cutting its membership from 50 to 14. "We wanted to take this membership issue... out of the lexicon," says Frank Maisano, a GCC spokesman. "It's a distraction from the real debate."

Mia Walton, a spokeswoman for GM, says that GM's move reflects a desire "to take a global, holistic, more 'GM' approach to the issue" of climate change. "The Global Climate Coalition has tended to be a North American or US-focused entity," she says. GM has plants in 53 countries and sells products in over 200. Walton adds that GM's opposition to the Kyoto Protocol still stands. Like the GCC, the company opposes the treaty's binding targets and timetables, and objects to the fact that it does not require participation on the same terms by all countries.

Some observers say that the corporations' withdrawals may simply reflect the lower prospects of ratification of the Kyoto Protocol by the US Senate. And with little prospect of ratification imminent, it is said, the companies can afford to polish their public images by withdrawing from GCC.

Whatever the motivation, the move appears to have weakened the GCC. Other major companies have left the lobby group in the past. In 1998, the British/Dutch oil company Shell withdrew, and in 1996, BP did the same (see *Nature* 392, 856; 1998 and 383, 470; 1996.)

Meredith Wadman

Could AIDS treatments slip through patents loophole?

Washington

A controversial patent issued last month on the gene for a key HIV target protein might not have been approved under recently revised patent guidelines.

These may open a loophole allowing companies to develop anti-AIDS drugs using the target protein, CCR5, without paying licence fees to the patent-holder, Human Genome Sciences (HGS) Inc. of Rockville, Maryland. But it is not yet clear whether the new guidelines — reflecting Washington's shifting perceptions about the use of patent law in this area — would apply, since the application was probably examined before they took effect.

The patent award continues to upset the researchers who characterized CCR5 in 1996 and showed for the first time that it is required for efficient HIV replication. They believe that they — and not HGS — should hold the rights to the chemokine receptor as a potential AIDS drug target.

When HGS filed the patent on the gene for CCR5 in 1994, the company had no idea of the gene's role in HIV, and was expecting instead to exploit the patent primarily in the development of anti-inflammatory therapy.

Gene patent rules have tightened up considerably since then. Applicants must now demonstrate "specific and substantial utility" (see *Nature* 403, 3, 2000).

William Haseltine, HGS's chief executive officer, says that the company's CCR5 patent meets all the new criteria for exploitation in the anti-inflammatory area, which is specifically mentioned in the patent. But he admits that AIDS therapeutics is "another question".

HGS expects that prospective drug developers wishing to exploit CCR5 to treat AIDS

will have to pay the company a licence fee. But some companies may not follow this practice.

Schering-Plough has already shrugged off HGS's intellectual property claim, having started developing a small CCR5-binding molecule before HGS's patent was approved. "Right now, their patent has no bearing on our compound," says spokesman Robert Consalvo. The company does subscribe to HGS's gene and protein databases. If it used that information to modify its drug, Schering-Plough would have to pay HGS royalties if the drug reached the market.

Haseltine points out that the scientists can make their own patent claims. Any additional patent approved by the US Patent and Trademark Office could complement the HGS patent, he says, hinting that drug developers might then have to 'cross-license', or pay two separate entities to work on one drug target.

Ned Landau, now an associate professor at the Salk Institute, was a member of one of five groups who raced to publish papers characterizing CCR5 in 1996. He and other researchers are also upset that HGS filed its patent primarily on the basis of computational searches, rather than experimental work.

"We did a lot of work," Haseltine says in reply, pointing out that computational searches involve making a large collection of human messenger RNA, conducting homology searches for existing genes, then focusing on specific families of proteins of potential interest to pharmaceutical partners.

Haseltine says that the company will not enforce its patent claim on academic researchers. HGS holds 114 gene patents, and has filed about 2,750 more.

Paul Smaglik

Israeli R&D grants 'too restrictive'

Jerusalem

The Israeli government should attach fewer strings to research and development grants, says a new report.

Tel Aviv University economist Manuel Trajtenberg submitted his report last week to the Office of the Chief Scientist of the Ministry of Industry and Commerce. Small and medium-sized high-tech firms should be given preference over large ones, he says, and successful companies should not have to reimburse the government. He suggests scrapping the requirement for companies receiving government R&D funds to produce any resulting goods in Israel.

Trajtenberg also recommends expansion

of the chief scientist's 'Magnet' programme funding cooperation between private companies and academic institutions. This would, he believes, counter the trend of small start-ups selling their technology abroad.

The report has been welcomed in Israel's high-technology community, though there is some scepticism about whether its advice is likely to be implemented.

Yossi Sela, managing partner at venture capital group the Gemini Fund, notes that many small companies don't even apply for government funds because of the restrictions and conditions. He points out that Israel's strength is not in cheap production but rather in the technology itself.

Haim Watzman

Northern lobby attacks UK synchrotron siting

London

So, the largest single investment in science infrastructure in Britain in two decades is to be made at the Rutherford Appleton Laboratory (RAL) near Oxford. But the decision on the site of a £200 million (US\$313 million) synchrotron (see *Nature* 404, 221; 2000) has sparked bitter protests from supporters of the rival Daresbury Laboratory near Manchester.

The synchrotron is a partnership between the British and French governments and the Wellcome Trust. Its powerful X-ray beamlines will be used to determine the structure of biological molecules and advanced materials. "The most important thing is that a decision to go ahead has been made," says John Finney, of University College London, a physicist who also studies biological molecules. Already, a year has been spent wrangling over the synchrotron's location.

Labour Members of Parliament from neighbouring northwestern constituencies lobbied hard for Daresbury. They had the ear of the trade and industry secretary, Stephen Byers. But the Wellcome Trust, and the Office of Science and Technology, which Byers oversees, argued for RAL. The decision was finally taken at cabinet level with Prime Minister Tony Blair directly involved.

RAL's backers argue that there are synergies to be gained from housing the synchrotron next to other facilities at RAL, such as the ISIS neutron source. Building up the RAL campus might also make it more attractive as a potential site for the planned European Spallation Source, another neutron facility.

But Daresbury's supporters complain of a lack of transparency over the decision. "This is not a spat, this is war," says Ian Gibson, MP for Norwich North and formerly Dean of Biological Sciences at the University of East Anglia. Gibson is also a member of the House of Commons Science and Technology Committee, which last year argued that there was no clear scientific case for favouring one site over the other.

Many MPs are angry about the possibility that the decision will widen the economic gap between the prosperous south and the depressed north. "The distribution of science excellence should be closely allied to regional development policies," argues Gibson.

The large and vocal lobby behind Daresbury, which includes 70 Labour MPs, five local universities, and two Nobel Prize winners, is particularly annoyed with the Wellcome Trust. Including running costs over 25 years, the synchrotron will cost £550m. The trust will contribute £110m, and the government says that the trust's preference for RAL was the deciding factor in the siting decision.

Helen Southworth, MP for Warrington South, complains that trust officials refused



Southern comfort: the Rutherford Appleton Laboratory, victor in synchrotron decision.

to meet a delegation of MPs. "I find that incomprehensible," she says. A spokeswoman denied that Wellcome Trust director Mike Dexter had refused to meet the MPs, but said he had been unable to find time to do so.

Staff at Daresbury, which is home to

Britain's existing national synchrotron, are similarly enraged. "The people who have got this work off the ground are at Daresbury; then it is stolen at the last minute and that really hurts," says a union representative.

Some users fear that vital expertise will be lost as technical staff take other jobs, rather than transfer to RAL. They are also worried about a possible gap in the availability of experimental facilities. "I suspect that many beamlines, including the one I am using now, will stop functioning properly because of lack of expertise well before the new machine becomes available," says Michael Ferenczi, a biophysicist at the National Institute for Medical Research in London.

Meanwhile, Southworth and other pro-Daresbury MPs are vowing not to give up, and met the Prime Minister to press their case earlier this week. Many future users of the synchrotron are sympathetic with their fight, but there are fears that a prolonged row could delay construction work. This is due to start next year, and planning permission still has to be obtained. "Now the community has to try to rebuild across the serious rifts that have been induced by this whole affair," says Finney.

Natasha Loder

French minister feels the heat over Soleil

Paris

Arguments used by French science minister Claude Allègre in favour of cooperating with Britain on a new synchrotron rather than building a national source in France have come under fire in a parliamentary report. The report, from the Office of Evaluation of Scientific and Technological Choices, urges that France immediately explore building an X-ray source on French soil.

Allègre had justified his decision to join the British government and the Wellcome Trust in building the synchrotron ring, Diamond, primarily on the grounds that a national synchrotron source built in France would be an economic burden, and that he was keen to see all 'big science' facilities built through European-level cooperation.

But the parliamentary document, released last Thursday, argues on the basis of evidence from more than 40 witnesses from throughout Europe that a national source would be the most

economical way of meeting the needs of French researchers. It says that the total costs to France per beamline on Diamond would be at least 50 per cent higher than on a national facility.

The report also criticizes Allègre's suggestion that France later rent beamlines in Switzerland and Germany, arguing that these facilities are already oversubscribed. "With a key technology such as synchrotron radiation, it would be harmful for France to adopt a short-term vision, probably reducing expenses very little, and certainly compromising the long-term potential of French research," states the document, which was written by deputy Christian Cuvilliez (Partie Communiste, Seine-Maritime) and Senator René Trégouët (RPR, Rhône).

When Allègre announced the French partnership in Diamond last summer, he also signalled the end of Soleil, a proposal ten years in the making to replace France's only national synchrotron facilities

at the Laboratoire pour l'Utilisation du Rayonnement Electromagnétique near Paris.

"The exclusive resort to international cooperation to satisfy national needs would appear somewhat paradoxical when France finds itself equipped with competitive resources...[and] the notion of scientific competition is more than ever present in the strategies of developed countries," states the report.

Cuvilliez and Trégouët support French participation in Diamond as an extra investment in synchrotron radiation to benefit disciplines such as biology. But during a press conference last Thursday, Cuvilliez called for the project to build Soleil to be relaunched.

The report is unlikely to have an immediate impact on the government's thinking. But it could be politically damaging to Allègre, who is already facing strong criticism from the ranks of his own Socialist Party for his handling of educational reforms.

Heather McCabe

US/UK statement on genome data prompts debate on 'free access'

The genomics company Celera is about to release the terms of access to its human genome data. Many will be reassured. But concerns remain.

Celera Genomics has promised to make the terms of access to its human genome sequence database publicly available within the next few days. The statement ends a week of fierce public and political debate over who should benefit from the use of genetic information.

A Celera spokesman says it will not place any constraints on the way the data is used by academic researchers, apart from not allowing it to be redistributed. Also it will not be demanding any 'reach through' rights — a response to accusations that it aimed to exert excessive control over the work of other researchers based on its data.

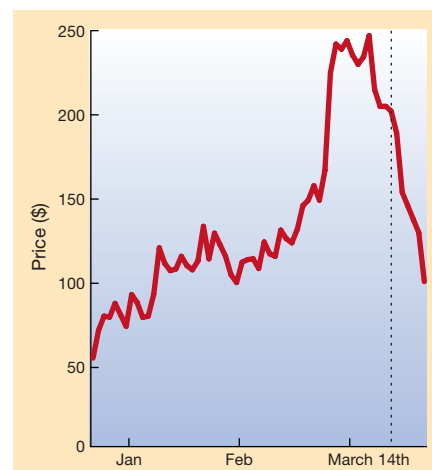
This move comes hard on the heels of a joint statement by US President Bill Clinton and British Prime Minister Tony Blair, applauding "the decision by scientists working on the Human Genome Project to release raw fundamental information about the human DNA sequence and its variants

rapidly into the public domain" and recommending other scientists around the world to adopt the same policy.

Their implicit target was Craig Venter's Celera Genomics. Paul Gilman, Celera's director of policy planning, says that typical licence agreements will be made available later this week. The release is likely to end months of speculation over terms of access.

"There are, unquestionably, a variety of significant issues in our contract negotiations," says a senior official at one large US university that has been negotiating a licence with Celera. "However, these issues are being discussed in detail with Celera, and they are clearly striving to identify terms that will allow academic investigators to have access to their sequence information, while also protecting their own interests. I think our differences could soon be ironed out."

Gilman says anyone will be able to access Celera's suite of data, software tools, annota-



Roller coaster: Celera's shares have fallen sharply since last week's statement by Clinton and Blair.

Japan may place gene research on summit agenda

Tokyo

Japan is considering placing human genome research on the agenda for the upcoming Group of Eight (G8) summit to be held in Okinawa in July. Members of the ruling Liberal Democratic Party (LDP) last week expressed strong support for the joint statement released by Tony Blair and Bill Clinton calling for free access to raw sequence data (see above).

At a meeting on the life sciences last week hosted by LDP and attended by both scientists and politicians, Taro Nakayama, former foreign secretary and leader of LDP's science committee, called for the issues surrounding human genome research to be taken up at the G8 summit, and urged Japan to respond positively to the joint statement.

Keizo Obuchi, Japan's premier, is also said to be in support of this move to place the issues on the G8 agenda. He told the daily *Nikkei* newspaper that "it would be great if different countries decide to take a united approach [on the release of human genome data]".

But the science-related ministries are wary of becoming involved. "Politicians are all for the free access of raw human genome data, but bureaucrats are extremely reluctant [to make a commitment to this]," says one leading genome researcher present at the LDP meeting.

Japan was approached by the US and British governments last year to take part in the joint declaration. A tripartite agreement was never reached because the ministries — particularly the powerful Ministry of International Trade and Industry (MITI) — were concerned about references to "release of variants [of human DNA sequence]", in a draft of the declaration, according to a source close to the Japanese government.

Japan is currently carrying out its own initiative to identify single nucleotide polymorphisms (SNPs) in the Japanese population, as well as a full-length complementary DNA (cDNA) project aimed at sequencing 30,000 cDNA clones. Both SNPs and cDNA are considered as variants of human DNA sequence.

Asako Saegusa

tion, and supercomputing power through annual subscriptions ranging from \$5000 to \$20,000, depending on whether they are in not-for-profit or commercial organizations.

Non-subscribers will be able to access the basic consensus sequence at Celera's website or on a DVD disc "without restriction", he says. "They can publish data and file for intellectual property protection on their discoveries. We will not seek any 'reach through' shared royalties or fees on any discoveries they make." He adds that both subscribers and non-subscribers will be allowed to use the data to develop chips and claim intellectual property on it — "except for our competitors using it in competition with us".

At the same time, Gilman admits that Celera will only be able to achieve a full sequence of the human genome by this summer if it takes much of its data from the public human genome project, and that anyone could wait just a few months and have free use of an equivalent amount of data assembled single-handedly by the public project. People will go to Celera not for raw data but for sophisticated computing tools, he argues.

The moves have been welcomed as important concessions by concerned scientists. But several still contest the fact that scientists will not be able to redistribute the data. This is important for annotation, says one, when researchers annotate a sequence that is then made available to others.

Gilman says Celera "would not want the public project to take our data and place it in

GenBank for other commercial competitors to reuse it". But one leading genome scientist says, "People should be free to circulate complete copies of the database, with additions of their own, if they want to; that makes for healthy bioinformatics. If they are not allowed to do this, the data is not being released according to GenBank standards."

Gilman dismisses this. He claims that the public project rejected an offer by Celera to put the data in GenBank, provided that those accessing it accepted that any use of the data



Venter: ending speculation.

would be freely accessible and not used to compete with Celera. "They said the terms of their international agreements are such that they can put no such limitation on the data."

The public project has been studying a similar 'click-wrap' contract, but this time intended to ensure that all those taking data — including Celera — agree to keep it open to all other potential users.

One burning question is whether leading scientific journals will agree to publish Celera's human genome, or whether they will prefer to publish that of the public project. Flood Bloom, for example, editor in chief of *Science*, is expected to publish a statement on this in tomorrow's issue of his journal.

At present, *Science's* policy is that "archival data sets (such as sequence and structural data) must be deposited with the appropriate data bank and the identifier code should be sent to *Science* for inclusion in the published manuscript," adding that the coordinates of this deposition "must be released at the time of publication".

Nature's policy is to require sequence data to be deposited in GenBank or a database of equivalent unrestricted accessibility, with an accession reference included in the publication. "Clearly we need to keep abreast of the changing landscape of databases, and the increasing involvement of private interests," says Philip Campbell, the editor of *Nature*. "It would be absurd to be fundamentally opposed to private database ownership, but the confidence of researchers and the public's stake in the content of the human genome are both of paramount concern." (See page 317.)

But some researchers are concerned that *Science's* policy does not explicitly require open access to data. Bloom told *Nature* that the magazine "will continue to advocate free access to nucleotide sequence data. Public databases are facing a challenge from the clash between the needs of two cultures — academia and industry. We at *Science* urge open, constructive dialogue between all parties so that unrestricted access to information can be assured, while still allowing enough protections that the biotechnology industry can flourish."

Declan Butler

'The human genome itself must be freely available to all humankind'

Bruce Alberts, president of the US National Academy of Sciences, and Sir Aaron Klug, president of the Royal Society of London.

"Through a public-private initiative of international proportions, science is about to provide the world with one of the most significant intellectual achievements of all time — the complete sequence of the human genome — which will provide a detailed map all of the genes in each cell that provide the blueprint for human life.

The implications are enormous. Last week, US President Clinton and UK Prime Minister Blair issued a joint statement applauding the effort. They emphasized the decision by scientists working on the Human Genome Project to release rapidly all of the information emerging from the project into the public domain, and urged scientists around the world to adopt the same approach.

We commend this powerful statement. But we also know that action by the world's scientists is not enough. More attention must be given to striking the appropriate balance between public and commercial interests.

Determining the sequence of the genome is similar to completing the list of the chemical elements: it tells us about the basic components, but not about how they behave in combination. In other words, it gets us to the starting line for a massive increase in understanding, but does nothing by itself to provide us with that understanding.

With the completed genome sequence, we will have all the instructions for making the 50,000 or more proteins present in the human body. But we will not know what each protein is for; still less about how the thousands of genes work together to produce and maintain the human body. And we will not know how the expression of a gene or genes can go wrong in the course of a disease.

In short, a huge amount of work will remain to be done. This will require effort in both the public and private sectors for generations to come. It is likely to lead to numerous breakthroughs in health care, benefiting all sectors of society. It is also likely to lead to numerous legitimate opportunities for creating wealth, underpinned by patent protection for the inventions and innovations of the individuals and companies involved.

What type of patent protection makes sense? Patent laws are based on the principle that public disclosure of a valuable new invention through the patenting process

should be encouraged and rewarded, allowing others to use and build on the invention to create additional benefits to society. In return, the inventor is rewarded with monopoly rights over the use of the invention for a limited period of time.

It is critical that the benefits to the public be at least reasonably commensurate to this reward. Given the enormous potential of the human genome sequence, the granting of broad monopoly patent rights to any portion of it should be regarded as extraordinary — and occur only when new inventions are likely to confer benefits of comparable significance for humankind.

It is a trivial matter today — using a computer search of public databases — to use DNA sequences to identify new genes with particular types of biochemical functions. In our opinion, such a discovery should not be rewarded with a broad patent for future therapies or diagnostics using these genes when the actual applications are merely being guessed at.

The intention of some university and commercial interests to patent the DNA sequences themselves, thereby staking claim to large numbers of human genes without necessarily having a full understanding of their functioning, strikes us as contrary to the essence of patent law.

Those who would patent DNA sequences without real knowledge of their utility are staking claims not only to what little they know at present, but also to everything that might later be discovered about the genes and proteins associated with the sequence. They are, in effect, laying claim to a function that is not yet known or a use that does not yet exist. This may be in current shareholders' interests. But it does not serve society well.

Scientists are at the very early stages of this work: knowledge of the human genetic sequence is only the first step. The next will be to understand how the tens of thousands of genes that make up the genome work together to create the machinery that operates the human body.

This will be a huge scientific undertaking. For it to work effectively and bring widespread benefit as quickly as possible, it is vital that all researchers have access to the full genome without charge or other impediment. The human genome itself must be freely available to all humankind." ■

US publishes rules on how data must be made public

Washington After a prolonged and fractious political debate, federally funded scientists in the US are now legally required to make certain data public under the Freedom of Information Act (FOIA).

The US government made the new requirement official last week. Fifteen government agencies published in the Federal Register the requirement that, from 17 April, government-supported scientists must “provide, within a reasonable time, [requested] research data so that they can be made available to the public”.

The change in FOIA came as a result of an amendment made by Senator Richard Shelby (Republican, Alabama) to a 1999 appropriations bill. He required government agencies “to ensure that all data produced under an award will be made available to the public” through FOIA.

The Office of Management and Budget (OMB) last November interpreted this language to require only the divulging of data used by the government in developing any action “that has the force and effect of law” (see *Nature* **401**, 732; 1999). The agencies are continuing to accept comments

on last week’s announcement until 15 May. The full text of the OMB ruling is available at <http://www.whitehouse.gov/omb>.

US physician fired over embryo scandal

San Francisco The University of California fired a tenured professor, Sergio Stone, last week after a reproductive medicine scandal. The offences happened five years ago at its Irvine campus, where eggs and embryos were misappropriated from patients (see *Nature* **376**, 546; 1995). Stone was found to have engaged in research misconduct involving patient studies, and had diverted substantial sums from university accounts.

A UC Irvine faculty committee had recommended demotion. But the UC Board of Regents voted last week to fire Stone, despite an emotional appeal by the physician, who said he was a scapegoat for the more egregious offences of two colleagues who had fled the United States after being indicted.

DFG takes on more and younger referees

Munich In a move to increase the transparency of its much-criticized review system, the Deutsche Forschungsgemeinschaft (DFG), Germany’s main grant-giving agency for

university research, has published the names of its 650 newly elected honorary referees on the Internet.

The DFG decided last year to increase the number of its referees by a quarter — and to encourage the enlistment of younger scientists, and more women — after complaints about its conservative funding policy and lengthy review procedures (see also *Nature* **404**, 217; 2000). The efforts appear to have borne fruit. Fifty women were elected to the latest referees’ list, raising the proportion of female participants to nearly eight per cent. The average age of referees decreased by 2.5 years to 53.2 years (http://www.dfg.de/organisation/fachgutachter/index_fr.html).

IMAGE launches series of space Explorer missions

Washington The US space agency NASA plans to launch a satellite to take regular global images of the swarm of charged particles surrounding the Earth on 25 March. Rather than use particle detectors that sample only the region they pass through, the \$154 million Imager for Magnetopause-to-Aurora Global Exploration (IMAGE) carries a suite of instruments for neutral atom, ultraviolet, and radio imaging of the entire inner magnetosphere.

The spacecraft is the first of NASA’s new

Medium-class Explorer (MIDEX) missions. Its orbit will change over the course of its two-year mission, which aims to view the Earth's plasma environment from different perspectives and build up a three-dimensional map. The principal investigator for IMAGE is James Burch of the Southwest Research Institute in San Antonio, Texas.

France to set up BSE testing programme

Paris France is to set up a nation-wide testing programme to gather data on the spread of bovine spongiform encephalopathy (BSE), its health and agriculture ministers announced on Saturday. The country came under fire last month when a European Union report found that traces of meat and bone meal continue to be found in French animal feed, despite a ban on this practice introduced in 1996.

A new case of BSE was reported by a farming collective last Friday, following the eighth case reported by the ministry of agriculture in February. France plans to administer the test on 40,000 cows that died or were killed after being suspected of being infected with BSE. According to the two ministers, the new plans go beyond the actions now under discussion by the European Union.

Britain starts GM crop trials despite gene-flow protests

London The British government last week published the locations of 30 new trial sites for genetically modified (GM) crops, and says that up to 80 more could be planted this year. The news follows reports that the Government had been having trouble finding enough farmers willing to participate in the trials, after environmental protesters destroyed a number of GM crop sites last year.

A consortium of researchers, led by the Natural Environment Research Council's Institute of Terrestrial Ecology, will run the trials. These will study biodiversity on GM and non-GM trial sites, as well as the potential and actual gene flow through cross-pollination from the GM to non-GM crops. In 1999, preliminary studies were carried out on nine sites.

Free Internet access to veteran satellite's data

Munich Nearly twenty years' worth of data gathered by the International Ultraviolet Explorer (IUE), a satellite launched in 1978 by the US, European and British space agencies, is being made freely available through the Internet. When IUE was switched off in 1996, life expectancy had been exceeded by

some fourteen years, and its output has become the most heavily used astronomical database in existence.

The European Space Agency (ESA) has created a new, decentralized distribution system, INES, which allows the IUE archive to be accessed freely and quickly via the Internet. The wealth of information from the more than 110,000 ultraviolet spectra gathered by the mission still promises to become a "mine of discoveries for future astronomers", says Willem Wamsteker, an ESA IUE project scientist.

Canadian science takes to the sky after a 30-year gap

Montreal Canada's first scientific satellite since 1971 is due to be launched in two years' time. The satellite, SCISAT-1, will study chemical processes involved in ozone depletion in the atmosphere, particularly over Canada and the Arctic.

Two contracts totalling Can\$13 million (US\$8.9 million) have been announced by Canada's Industry Minister, John Manley. They have been awarded to Bristol Aerospace of Winnipeg, a company that has been building atmospheric sounding rockets and payloads for rocket and space shuttle missions for more than 30 years. SCISAT-1 will be Bristol's first satellite.

To create generalists, teach students how to learn by themselves

Sir—In his Millennium Essay, Frederick Seitz¹ makes a compelling case for the importance of generalists — individuals with broad vision — in all cultural fields, and he laments that their ranks are thinning. I subscribe entirely to Seitz's assessment, yet I disagree with his analysis of what causes this trend and with his suggested remedy: reforming elementary and secondary education.

It has become almost a cliché in the United States and Europe to blame schools for society's increasing cultural ignorance. Yet many schools have struggled for years to maintain a diverse curriculum, despite budgetary constraints and conflicting societal demands.

Even so, a school's ability to affect students' cultural breadth is very limited. Children spend most of their time with family and friends, not at school. Parents need to provide daily evidence that they value diverse cultural pursuits: if they let their children devote an inordinate amount of time to a narrow range of activities, schools cannot be expected to produce people with differing values.

Provided parents have planted the seed of cultural and intellectual curiosity in their children, there is little that colleges and universities can do to kill it, but they can nurture it to a lesser or greater extent. Seitz considers that university teaching staff have virtually no room to manoeuvre against the strong internal and external forces that promote specialization. Chief among these, according to Seitz, are "conditions of intense competition", which leave little time for scholars to cultivate new, diverse interests.

Examples abound, however, of scholars such as Einstein, Jaynes and Mandelbrot who gained prominence and significant competitive advantages in their field precisely because they were able to borrow concepts, tools and methods from disciplines far from their own.

The growing complexity of most fields of research, which Seitz also mentions as an incentive to specialize, has to be put in historical perspective. A century and a half ago, scholars in North America and Europe were expected to be well versed in all branches of knowledge, and to remain so during their whole lives.

Admittedly, most fields of research have become much more complex since then, but the level of specialization of most scholars has increased incommensurably. I believe that colleges and universities are responsible for this, not because of

pressures on them but because of their approach to learning.

By the second half of the nineteenth century, undergraduate education was in many respects similar to what we have now and, except in a handful of English institutions such as Oxford and Cambridge universities, did little to equip students with the skills necessary to keep learning new disciplines and updating their knowledge. On the other hand, an old tradition of "self-study" in society at large provided them with these skills². Early graduate programmes in Germany and, for a time, at Johns Hopkins³ emulated this in their system of "self education under guidance".

In the early part of the twentieth century, however, universities eradicated self-directed learning from the educational landscape. They successfully promoted the idea that quality learning can occur only in the physical or, more recently, 'virtual' presence of a teacher. With rare exceptions^{4,5}, colleges and universities nowadays do not prepare students to keep learning on their own after they leave college. As a result, many graduates (including many who become academics) find that their lack of skills renders learning inefficient and slow. The best they can do is to deviate as little as possible from the narrow trajectory on which they were placed during their initial training, or retraining in the private sector.

To alleviate the resulting cultural and intellectual atrophy, colleges and universities must rethink their attitude towards self-directed learning, and implement imaginative ways to foster it effectively.

Philippe Baveye

Laboratory of Environmental Geophysics, Bradfield Hall, Cornell University, Ithaca, New York 14853, USA and School of Education, University of Paris X, 92001 Nanterre, France

1. Seitz, F. *Nature* **403**, 483 (2000).
2. Kett, J. F. *The Pursuit of Knowledge under Difficulties* (Stanford Univ. Press, Stanford, 1994).
3. Ludmerer, K. M. *Learning to Heal: The Development of American Medical Education* (Johns Hopkins Univ. Press, Baltimore, 1985).
4. Baveye, P. in *Current Developments in Self-directed Learning* (eds H. B. Long et al.) 157–175 (College of Education, University of Oklahoma, Norman, 1996).
5. Candy, P. C., Crebert, G. & O'Leary, J. *Developing Lifelong Learners through Undergraduate Education* (Australian Government Publishing Service, Canberra, 1994).

We have touched the dust from dying stars

Sir—Adam Burrows's superbly eloquent review article¹ about supernova explosions omits one aspect of supernovae that is so stunning that it adds to the context of his essay. Namely, that the human race holds solid samples of supernovae in its hands, and studies those samples in terrestrial laboratories.

Mostly micrometre-sized balls of SiC and graphite, they contained at the time of their condensation such staggering concentrations of ⁴¹Ca, ⁴⁴Ti and ²⁶Al that the huge daughter abundances unequivocally declare these grains to be condensates within the interiors of their expanding supernova hosts². Very non-solar isotopic ratios of the stable elements C, Si and N within them confirm this diagnosis. These solid grains are local pieces of diverse supernova interiors, before any mixing with circumstellar matter. As part of the interstellar dust, they found their way into meteorites when those bodies grew at the birth of the Sun. They illuminate unforeseen details of supernova nucleosynthesis and supernova hydrodynamics.

Donald D. Clayton

Department of Physics and Astronomy, Clemson University, Clemson SC 29634-0978, USA

1. Burrows, A. *Nature* **403**, 727 (2000).
2. Clayton, D. D., Amari, S. & Zinner, E. *Astrophys. Space Sci.* **251**, 355 (1997).

Confusion over cash for Indian biotech centre...

Sir—Despite your News report "Funding crisis for Indian biotech centre" (*Nature* **403**, 694; 2000), we have not halted recruitment at the Indian branch of the International Centre for Genetic Engineering and Biotechnology (ICGEB). Several scientists have recently joined the centre and none has left. Our financial support from the Indian government and from Italy is continuing without reduction. Furthermore, the Indian government has recently provided additional support for several new programmes at our centre, particularly the malaria vaccine project.

V. S. Chauhan

Director, ICGEB, PO Box 10504, Aruna Asaf Ali Marg, New Delhi 110 067, India

... though funding states have promised to pay up

Sir—It is not true to say that the member states are refusing to pay their contribution to the ICGEB; on the contrary, at the latest meeting of our board of governors, all the delegations expressed their firm intention to fulfil their statutory financial obligation according to the agreed scale of contributions.

Furthermore, the New Delhi component of ICGEB is not funded by "grants" from Trieste: ICGEB is a single entity and its two components in Trieste and New Delhi are both funded from the ICGEB regular budget, an essential

element of which comes from the Indian government. Finally, there is no restriction in the centre's budget: the two host governments (Italy and India) both intend to increase their contributions, and we are confident that the other member states will fulfil their obligations. If any minor financial restrictions are made, they would affect equally both the Italian and the Indian components.

Arturo Falaschi

Director, ICGEB, Padriciano 99,
34012 Trieste, Italy

K. S. Jayaraman replies — I stand by my story. When I spoke to him, Dr Chauhan confirmed my information that the last scientist to join the centre at senior level did so four years previously. The cuts in the centre's grants (10 per cent now and a total of 25 per cent by 2001) were also confirmed by Dr Chauhan during our conversation.

The lasting value of Mitchell's mechanisms

Sir — As Leslie Orgel¹ and Bo Malmström² have valuably reminded us, Peter Mitchell opened up new vistas in bioenergetics when he formulated his chemiosmotic theory. The theory itself was a special case of his more general view on vectorial metabolism, a concept which little influenced his contemporaries although it formed a critically important part of his own intellectual odyssey. The chemiosmotic theory, initially formulated without experimental support, in time provided a framework for the development of our modern understanding of bioenergetics.

Was Mitchell never right on the mechanistic level? He certainly came up with mechanisms in large numbers. When I visited him late in his life he showed me several of them, covering various questions. There were yet more piled up on the window ledge. His private papers include many more. It is true that most of these ingenious schemes led nowhere, including those covering the involvement of protons in the ATPase and in the cytochrome oxidase, as well as some curious schemes for metabolite and ion transport. Indeed, Mitchell was allergic to systems that depended on processes which at the time could not be defined precisely at the molecular level, such as the conformational processes that he regarded as 'black box' biochemistry because they relied on an unspecified and unknown mechanism.

However, one mechanism that came to him in the middle of a sleepless night has been an outstanding success. This is the Q cycle³ for the transport of protons across

the inner mitochondrial membrane by the ubiquinone-cytochrome *c* reductase complex. This surely gave major support to the chemiosmotic theory in a period when the validity of the theory was itself still under debate. Perhaps this one achievement, together with the chemiosmotic theory with all its implications, justifies so many apparently fruitless mechanisms.

John Prebble

School of Biological Sciences, Royal Holloway
College, University of London, Egham Hill, Egham,
Surrey TW20 0EX, UK

1. Orgel, L. *Nature* **402**, 17 (1999).
2. Malmström, B. *Nature* **403**, 356 (2000).
3. Mitchell, P. *FEBS Lett.* **56**, 1–6 (1975).

People must be judged in the context of their time

Sir — Alison Abbott reports on the Max Planck Society's recent efforts to shed light on its history, as the Kaiser Wilhelm Society, during the Third Reich (*Nature* **403**, 474–475; 2000). This attempt is all the more important as investigations in the past encountered some reserve, whereas new information is now available and a new generation of science historians is examining the past, supposedly without any preconceptions.

However, the example of Hans Stubbe, first director of the Kaiser Wilhelm Institute (KWI) for Cultivated Plant Research (established near Vienna in 1943 and built up after the war at Gatersleben in East Germany) raises doubts as to the investigators' impartiality. Portraying Stubbe as "an opportunist, ruthlessly following any path that would help him to advance his research and career" and saying that he "collaborated with the SS to plunder valuable Russian collections of wild and cultivated plants" is unjustified.

The Russians themselves judged differently, since Stubbe received their backing to build the institute at Gatersleben and was highly honoured by the Soviet Union in spite of being an outspoken opponent of Lysenkoism — an intervention that saved East German genetics from the damage inflicted in other Eastern bloc countries. Contemporaries know that Stubbe protected several colleagues who faced difficulties with the Nazi or the East German system and enabled them to continue their work. He had himself been expelled from the KWI in 1936 for political reasons.

We can only beg those historians who never experienced a totalitarian regime to avoid hasty moral judgements and not to base their conclusions on what Jens Jessen has described in the German newspaper *Die Zeit* as "moral maximalism", without any feeling for the unavoidable

entanglement in guilt and the tragedy always involved in making any significant decision within a totalitarian system. We need a careful documentation of all the historical facts, but we can also ask that people be judged within the historical context in which they lived and worked.

Ulrich Wobus, Ingo Schubert

Institute of Plant Genetics and Crop Plant Research,
D-06466 Gatersleben, Germany

Heim replies — Wobus and Schubert say it is "unfair" and "unjustified" to claim that Hans Stubbe collaborated with the SS to plunder valuable Russian collections of wild and cultivated plants. But there is extensive historical evidence testifying to precisely these facts.

The extended documentary trail left by Stubbe in various German archives contains a great deal of evidence about his close connections to the SS, especially his collaboration on a broad level with the SS-Foundation Ahnenerbe, or 'Ancestral Heritage' — the Nazi research organization. A 1944 document bearing Stubbe's signature shows that members of his institute were carrying out "extensive research in the field of biological warfare". Documents of this time make it clear that he supported the 'rescuing' of valuable plant materials from Russian soil.

Alison Abbott pointed out correctly that Stubbe was not an "out-and-out Nazi"; he has to be seen in his ambivalence. He came under suspicion by the Gestapo, for example, while maintaining good relations with the SS. There is no evidence of anti-semitism in Stubbe's Nazi-era correspondence, although after the war he helped at least one former SS officer from Auschwitz get his name cleared.

It is this ambiguity which makes him interesting for studying the role of scientists during the Nazi era, because his combination of attitudes is probably more typical than the black-and-white cases.

Stubbe's merits as a geneticist in post-war East Germany and his recognition and appreciation by postwar Soviet authorities do not prove anything about his role during the Nazi period. Wobus and Schubert's reprimand of "historians who never experienced a totalitarian regime" recalls the outmoded idea that one cannot write good history without having lived through the same conditions as the people being described.

The more important lesson, I would argue, is that scientists concerned about history should be as willing to examine the bad as the good. Stubbe was indeed a complex figure who lived in a difficult time, but we do no service to the past or the present of science by ignoring its seamier sides.

Susanne Heim

Max Planck Institute for the History of Science,
Wilhelmstrasse 44, 10117 Berlin, Germany

The imperial slaughterhouse

Literature and the toll paid for the subjugation of the British Empire.

Romanticism and Colonial Disease

by Alan Bewell

Johns Hopkins University Press: 2000. 373 pp.
\$45, £35

Roy Porter

Living as we do in an era of the 'globalization of disease', threatened by new killer viruses jetting around the world, we have become more aware of the role of disease in shaping human history. And in no sphere is that more evident than in the blood-drenched chronicles of imperialism. Wherever Europeans went, they brought ghastly epidemics — smallpox, typhus, tuberculosis — to virgin populations entirely lacking resistance. The results were devastating. During the century following Christopher Columbus, in many zones of the New World up to 90 per cent of the indigenous population was wiped out — germs were far more lethal than guns in expediting the work of conquest. Over the course of the next three centuries, some 20 million slaves then had to be shipped to the Americas to fill the labour vacuum, and they in turn brought in their wake the characteristic African diseases of malaria and yellow fever.

It would be a mistake for us to assume that our post-colonial times are the first to appreciate the medical disasters wrought by empire-building. Compassionate contemporaries were under no illusions, as Alan Bewell stresses in his skilfully written, cross-disciplinary book. "Wherever the European has trod, death seems to pursue the aboriginal," rued the sensitive young Charles Darwin in his *Beagle* journal, shortly before the Tasmanian aborigines were driven into extinction, largely by white men's diseases.

Particularly sensitive to this melancholy plight were the English Romantics, sympathizing with victims and siding with the people against the powerful. The artist and poet William Blake never ceased to expose the plague of empire, and he found worthy successors in Percy Bysshe Shelley and Lord Byron, who himself fell victim to fever while fighting on the side of the Greeks against the curse of the Ottoman Empire; and even the increasingly reactionary Samuel Coleridge remained fervent in his hatred of the slave trade. Bewell brings out well the Romantics' dramatic vision of the toll taken by imperial conquest upon the health of all parties.

In poetry and prose alike, Romantic circles portrayed the hubris of the colonial slaughterhouse. The health of natives was undermined by settlers' diseases; slaves were brutalized, not least by syphilis passed on by masters' lusts; and Europeans fell like flies



Sickness from abroad: Father Thames presents his offspring, which includes the tropical disease cholera.

themselves, above all, those hard-living and hard-drinking soldiers sent to garrison the outposts of West Africa and the Caribbean. Out of 350 men sent to the Gold Coast in 1782, only seven were alive and on duty three years later. Despatched to quell a slave rebellion on St Domingue, 40,000 French soldiers died, allowing Haiti to come into existence, the first black state in the New World.

Like Coleridge in his *Rime of the Ancient Mariner*, William Wordsworth, too, was particularly effective in conveying the plight of 'tropical invalids'. Returning from duty overseas, pallid and haggard, their constitutions wrecked, ex-soldiers were treated as pariahs, disowned by all for bearing the stigma of colonial dirty work. It is no accident that it was Mary Shelley who, with the first cholera pandemic approaching, produced in *The Last Man* (1826) a futuristic novel about the eradication of the entire human race by a new killer plague emerging out of the East, nature's punishment for human vainglory.

The tropics also did rhetorical service as a mirror held up to home at a time of industrialization. The living conditions of the masses in 'darkest England' were routinely likened to those of Indian or African natives. "The room of a fever patient, in a small and heated apartment in London, with no perflation of fresh air," opined the sanitary reformer Thomas Southwood Smith, "is perfectly analogous to a stagnant pool in Ethiopia, full of the bodies of dead locusts. The poison generated in both cases is the same." Cholera introduced a

terrifying tropical disease into the very homes of the English, the retribution of the Orient. As in the subjugated empire, the disease was widely read as symptomatic of the filth and disorder of the masses, indeed, a manifestation of the 'democracy fever' that arose with the Reform Bill of 1832.

Perhaps galvanized by putrefaction and pauperism at home, Romantic attitudes to colonial disease were complex. On the one hand there was a pervasive fatalism. With their steamy swamps and rotting vegetation, the tropics were inherently pathogenic, according to the dominant miasmatic theories of disease aetiology set out in James Lind's *Essay on the Diseases Incidental to Europeans in Hot Climates* (1768) and elsewhere. Sensible diet, lifestyle and 'seasoning' (acclimatization) might reduce the risks, but torrid climes were inevitably white men's graves.

Yet Romanticism also possessed a Promethean streak. Through the application of science, energy and will, progress could transform the environment, drain swamps, fell forests, introduce sweet air, and clean up settlements; humanitarianism could make the 'heart of darkness' as hygienic as temperate parts. The Romantic dream of environmental transformation thus expressed a noble ideal, not dissimilar to the imperialist 'white man's burden'. As Bewell insists, these facts require us to abandon any simplistic vision of the Romantics as pure nature-lovers.

Readers familiar with this field will find

book reviews

that *Romanticism and Colonial Disease* sometimes reploughs furrows recently turned by others. *Jane Eyre* is recycled as an imperial novel, and the ambiguous attitudes of Thomas De Quincey towards the East (it brought cholera and the bitter-sweet opium to which he was addicted) are again discussed. And, strangely for a professional literary critic, Bewell is at his weakest when practising textual criticism, some of his readings being forced or far-fetched. We are told, for instance, that "And gathering swallows twitter in the skies", the evocative final line of John Keats' ode "To Autumn", incorporates a reference to the sore throat from which the poet had been suffering as a result of his encroaching tuberculosis, making swallowing difficult. Readers will surely find that interpretation hard to swallow.

Overall, however, this is a confident addition to the growing study of the history of environmental protest. It used to be said that if there was any justification of imperialism, it was the medical benefits it had conferred. As early as the Romantics, that could have been dismissed as special pleading. ■

Roy Porter is at the Wellcome Institute for the History of Medicine, 183 Euston Road, London NW1 2BE, UK.

A meteorologist's nightmare

Isaac's Storm

by Erik Larson

Fourth Estate/Crown: 1999. 366/384 pp.
£16.99/\$25

Howard B. Bluestein

My late paternal grandfather, Isaac Springer, described to me his experience of driving around fallen trees and downed power lines during the great New England hurricane of 1938. Although he did not know then that he was experiencing one of the most significant



Wave of destruction: a lithograph depicting the 20-foot storm surge that devastated Galveston.

North American storms of the century, he did realize that something different from the ordinary was occurring. Thirty-eight years earlier, on 8 September 1900, Isaac Cline, a meteorologist living in Galveston, Texas, faced a similar situation. This Isaac's storm, one of the greatest and most deadly hurricanes ever to strike the United States, was not forecast by the US Weather Bureau, largely because there were few observations over the ocean, and many forecasters incorrectly interpreted the data available. Consequently, thousands of people died in a storm whose winds reached perhaps as much as 150 miles an hour, and which produced a storm surge of some 20 feet, washing away houses and the people in them.

Cline, who survived, had not believed that a hurricane born in the Caribbean could strike the Texas Gulf coast; instead, he thought a Caribbean hurricane would curve

to the right as it progressed westward, so that it would make landfall along the east coast of the United States or along the Gulf coast east of Texas, or simply stay over the Atlantic.

Erik Larson's riveting *Isaac's Storm* is a fascinating account of the great Galveston hurricane, recreated around the central character of Cline. The book sets the historical stage for the hurricane and how the Weather Bureau mishandled its forecast. It is hard to conceive today, when we have excellent surveillance by satellites and aircraft both day and night, that even the existence or location of hurricanes over the ocean was not so long ago a matter of surprise at worst and of debate at best. We still have difficulty predicting the precise track of a hurricane, and whether or not it will intensify, because we do not have a map of the exact ocean temperatures under the storm, or details of the wind field throughout the troposphere in the vicinity of the storm. However, we do know that a hurricane is out there and therefore can try to guess what stretch of coastline it might strike by extrapolating its future track. As a result, watches and warnings alert citizens well in advance of hurricane landfall.

Why the hurricane was not forecast properly, however, was only in part a result of scientific uncertainty; the public was not warned mainly because of politics and personalities. The author carefully documents how government officials, possessed by vanity and intoxicated by power, could have allowed better forecasts despite the scientific uncertainties. The author makes a convincing case that the Galveston tragedy involved a mixture of scientific ignorance, managerial incompetence and bad luck.

New in paperback

Geometry Civilized: History, Culture and Technique

by J. L. Heilbron

Oxford University Press, £25, \$35

Fearful Symmetry: The Search for Beauty in Modern Physics

by A. Zee

Princeton University Press, \$14.95, £9.50

Shadows in the Sun: Travels to Landscapes of Spirit and Desire

by Wade Davis

Broadway, \$13

Noah's Flood: The New Scientific Discoveries about the Event that Changed History

by William Ryan & Walter Pitman

Simon & Schuster, \$14, £8.99

The User Illusion: Cutting Consciousness Down to Size

by Tor Nørretranders

Penguin, £10.99

Sex and Friendship in Baboons

by Barbara B. Smuts

Harvard University Press, \$19.95, £12.50

The narrative puts the reader in the hurricane itself; we feel anxiety as unusual swells reach the coastline far in advance of the hurricane. And we can virtually hear the deafening roar of the howling winds as victims are described as grabbing for any piece of floating debris to survive.

The effectiveness of Larson's story is enhanced by his careful scholarship. Extensive footnotes and sources are listed at the end of the book for the more serious reader, but are transparent to the recreational reader. Although Larson makes extensive use of artistic licence, the results are rarely not believable, even though I often wondered, "how does he know that?". Larson's extensive reading of correspondence, newspaper clippings and other sources at libraries, and his personal interviews, have been carefully pieced together, like the actions of a brilliant sleuth solving a crime or a mathematician solving a problem, to produce a highly realistic account of what must have happened.

A number of memorable anecdotes, some of which are not original to Larson's book, work well here. The thoughts and experiments of early scientists and 'meteorologists' are most interesting. The Earth's rotation has an important role in the dynamics of hurricanes through the Coriolis force. My favourite account is of Joseph Furtenbach's loaded-cannon experiment, in which he attempted to see whether the Earth really rotates about its axis. The fact that Furtenbach was not hit on the head by a cannonball proved that he was right, although one could certainly devise experiments to test a theory that are less risky for the scientist conducting them.

Additional explanations of the physics of hurricanes and other storms are simplified, so that the layperson can understand the essence of the mechanisms responsible for storm formation and movement. Larson weaves these explanations in with interesting historical narrative. However, I was not fully convinced of the importance of chaos in the formation of hurricanes, a phenomenon that the author referenced from a book by E. Zebrowski Jr. Whether the sea surface temperature exceeds a critical value or whether the vertical wind shear is below a critical threshold are probably much more important than the precise intensity or location of the incipient tropical cyclone in determining whether or not a hurricane actually forms. On the other hand, the track and intensity of a tropical cyclone are chaotic, that is, very sensitive to initial conditions.

The account of the rivalry between Cuban and US meteorologists is particularly charming (or embarrassing). During the peak of the hurricane season, the United States banned cablegrams from the Cuban government about the weather over its own lines, so that only information from the US

Weather Bureau was available, in spite of Cuban meteorologists' demonstrated skill in hurricane forecasting.

Larson proposed that the Cuban approach to forecasting was more romantic, relying less on measurements than on intuition, while the US Weather Bureau's approach was based more on data analysis and interpretation. I know of many theoreticians and experts on atmospheric dynamics and analysis who are not good forecasters; in addition, there are many forecasters who do not understand dynamics very well but who are, nevertheless, excellent forecasters with good intuition. Forecasting is partly science and partly art, because observations are subject to different interpretations.

Larson vividly describes the price paid for an inaccurate hurricane forecast or for ignoring a good one. There are gruesome accounts in the book of ships caught in hurricanes and of a train heading for Galveston during the hurricane. The ghastly descriptions of the aftermath of the Galveston hurricane range from the horrendous odours, and images of masses of corpses, to poisonous snakes and devastating building damage. This sobering account of a hurricane's effect eliminates any desire to experience the fury of a major hurricane while stranded on a low-lying strip of coastline.

I recommend this book highly to scientists, laypeople fascinated by storms, and social scientists. To prevent future disasters like the 1900 Galveston hurricane, advances in the prediction of hurricanes must be accompanied by efficient means of disseminating forecast information, and

by the correct response of people living along the coastline where hurricanes may strike. Perhaps this book will stimulate such action.

Howard B. Bluestein is at the School of Meteorology, University of Oklahoma, Norman, Oklahoma 73019, USA.

Palaeobiology in mammoth form

Encyclopedia of Paleontology, Volumes I & II

edited by Ronald Singer

Fitzroy Dearborn: 1999. 1,423 pp. £175, \$285

Michael J. Benton

Encyclopedias are all the rage. Major publishers clearly believe there is a market for massive, multi-volume tomes that tell the whole story about a subject. This might seem somewhat counterintuitive now that

most students go first to the World Wide Web if they want to find something out. Ronald Singer's new *Encyclopedia of Paleontology* is as big as any of the recent monster works, and it is certainly the largest encyclopaedia ever published on palaeontology. The critical test is whether this two-volume work gives a thorough account of



Life on the large side

The fragmentary fossil remains from the vast supercontinent of Gondwana now lie scattered across the five land masses of India, Africa, South America, Australia and Antarctica. *Wildlife of Gondwana: Dinosaurs & Other Vertebrates from the Ancient Supercontinent*

by Patricia Vickers-Rich and Thomas Hewitt Rich (Indiana University Press, \$59.95, £45) looks at the history of Gondwana's vertebrate fauna. Above, a graphic artist's giant goanna, *Megalania prisca*, threatens the dromornithid bird *Genyornis newtoni* in Pleistocene Australia.

modern palaeobiology, or whether, like the dinosaurs, it has had its day.

The list of contributors is impressive — some 200 research scientists from many countries, although mostly from North America. Some of these authors offer classic nuggets: Stefan Bengtson on the evolution of skeletons, Euan Clarkson on fossilized eyes of invertebrates, Philippe Janvier on craniates, Conrad Labandeira on fossil insects, Michael Ruse on various nineteenth-century palaeontologists, Dolf Seilacher on the Ediacaran biota, and Ian Tattersall on primates. Other articles are more workaday, and the editor clearly had difficulty plugging many of the gaps. Among historical figures, for example, we have William Buckland and Richard Owen, but not Gideon Mantell; Marcellin Boule and Eugène Dubois, but not Arthur Keith; Otto Schindewolf and Friedrich von Huene, but not Karl Zittel.

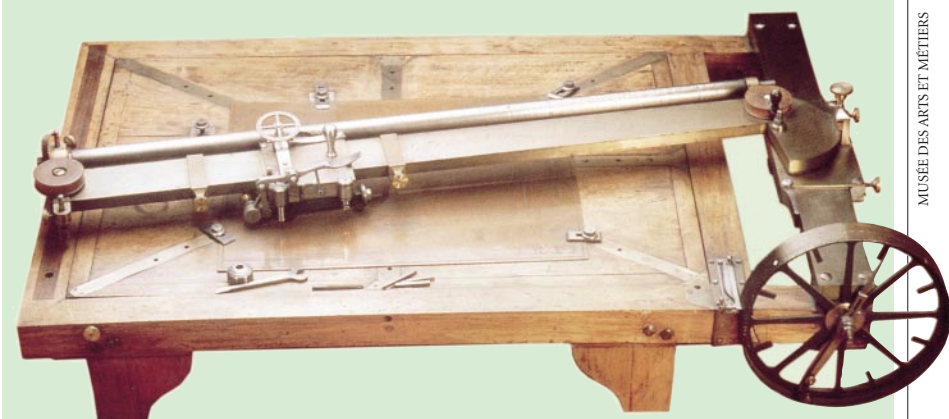
The biggest gap is in the area of systematics and analytical approaches in evolutionary theory. There is one article on systematics, and a series on evolution, variation, selection, extinction, and the like. But where is the fulsome representation of current palaeobiological research on mass extinctions, rates of evolution, diversification, morphometry, fossil-record quality, and cladistic methods? The only articles on 'events' are on the origin of life and terrestrialization. What about the five big mass extinctions, major changes in tectonics in the sea and on land, multiple origins of flight, the Mesozoic marine revolution, the spread of grasslands?

The accounts of specific groups of plants and animals are generally useful. You can look up a brief, up-to-date account of aardvarks, acanthodians, algae or annelids. Admittedly, though, such accounts are also available in current textbooks on invertebrate or vertebrate palaeontology, and on palaeobotany. The accounts of topics in palaeobiological theory and method are useful too, but, again, such information is also available in textbooks and review journals. Derek Briggs' and Peter Crowther's *Palaeobiology: A Synthesis* (Blackwell Scientific, 1990; new edition due 2000) offers much better coverage in the form of mini-essays designed both for professionals and for students. The inclusion of historical personages in Singer's *Encyclopedia of Paleontology* is a mixed blessing: many are excellent, pithy accounts, and provide key references, but many are missing, and some of the people included are perhaps not sufficiently important in the history of the subject to warrant a mention.

But the book is better than many encyclopaedias, and it contains some excellent nuggets. It will be a useful addition to university and museum libraries. ■

Michael J. Benton is in the Department of Earth Sciences, University of Bristol, Queens Road, Bristol BS8 1RJ, UK.

Science in culture



French fusions

The newly renovated Musée des Arts et Métiers in Paris.

Martin Kemp

There is no other collection quite like the Musée des Arts et Métiers in Paris. It looks, at first sight, like a museum of technology, yet it encompasses much more than this English term implies. It relies upon that special French fusion of arts, crafts, manufacturing, technology, engineering and science, laced with *philosophie* and backed by social reform.

Founded in 1794 by the revolutionary priest Abbé Henri Grégoire, the Conservatoire des Arts et Métiers was to serve as a repository for "tools and newly invented and perfected machines" in order to provide a source of didactic inspiration for French craftsmen. In particular, those inventive and visionary individuals "who can see further will make new connections, for all the arts have points of contact". The accumulated collection, housed since 1798 in the abbey of Saint-Martin-des-Champs (itself a wonder of medieval structural technology), had declined to a state of benign neglect when, some ten years ago, the current director, Dominique Ferriot, embarked on a heroic campaign of renovation.

Re-opened to the public this month, the old penumbral labyrinth inhabited by mysterious fossils of bygone technologies has been superseded in Andrea Bruno's bold design by bright airy spaces, enclosed by white walls, complemented by the gleam of exposed pipe-work and the shiny interplay of metal and glass. The wondrous machines, including such visionary marvels as Blaise Pascal's seventeenth-century calculator and Louis Blériot's precarious flying-machines, are literally shown in a new light, and elucidated by touch-screen animations. Grouped in seven domains — the scientific instrument, materials, construction, communication, energy, mechanics and transport — the exhibits range from early astrolabes to pioneer television receivers, and from Georges Buffon's compound mirror to sumptuous art nouveau glass by Emile Gallé.

The ingenious device that I have chosen to embody Grégoire's aspirations is an engraving

machine invented by Nicolas-Jacques Conté, whose name is immortalized in the art world by the types of pencil and crayon that still bear his name. Conté, trained as a painter, entered the technological world via ballooning, working on the large-scale production of hydrogen to fuel balloons from the breakdown of water. As a fertile inventor and committed pedagogue, he was an obvious candidate to be enlisted by the Abbé as one of the three original demonstrators in the Conservatoire.

In the great age of illustrated encyclopaedias and didactic publications aimed at the new bourgeoisie, any mechanical device that could short-cut the laborious business of reproductive engraving would meet a conspicuous need. Conté's beautiful machine, shown here in the version constructed by Gallet in 1803, was originally conceived to facilitate the production of the 900 plates for his 23-volume *La Description de l'Égypte*.

Precision-designed to cut lines regularly into copper engraving plates, his machine could be set up to produce wide varieties of straight and curved incisions, not least the laborious sets of parallel, curved and cross-hatched gouges needed to print areas of modulated tone in a meticulously controllable way. As a draftsman skilled of hand, Conté took pains to ensure that the engraving tool could incise the plate at varied pressures in order to emulate some aspects of the subtle 'touch' that gave such grace to the work of master engravers and etchers.

His engraving machine is a perfect manifestation of the eighteenth-century passion for automated devices, ranging from the restrictedly practical to the unashamedly entertaining (such as the recreational automatons of which the museum holds a supreme collection). Such marvels of artifice finally and decisively confirmed the triumph of the 'moderns' over the 'ancients'. In the Revolutionary era of Grégoire, Conté and his fellow inventors also played their role in constructing a new era in the social and economic machinery of society itself. ■

Martin Kemp is in the Department of the History of Art, University of Oxford, 59 George Street, Oxford OX1 2BE, UK.

Brains, courage and integrity

Gandhi and Sakharov set us an inspiring example for the twenty-first century

Tom Gehrels

Mohandas Karamchand Gandhi (1869–1948) and Andrei Dmitrievich Sakharov (1921–1989) were men of intelligence combined with integrity and the courage of their convictions, even at great personal cost. They were kindred souls, remarkably alike in their use of these attributes for the sake of peaceful progress.

Their original professions were very different. Sakharov's first assignment was to a munitions factory during the Soviet Union's war against Germany, when it was found that he could solve problems, increasingly more difficult, which led to his assignment to the nuclear weapons programme. At first he believed this work was necessary for a balance of power, but gradually his conscience drove him to consider the related issues of human rights and survival for the whole world. Gandhi began with studying for a law degree in England, but later, as a lawyer in South Africa, he found that his views on independence and human rights led him to develop a successful new career.

Gandhi and Sakharov were equally bright and courageous people, setting themselves goals to achieve, even at the cost of repeated hunger strikes. Gandhi tried desperately but unsuccessfully to prevent the partition of colonial India into the separate states of India and Pakistan. However, he saved many lives during the frictions between Hindus and Muslims. Sakharov was able to exert a powerful influence in the Soviet Union and on Premier Khrushchev, because he had originally worked within their system. A dissident of lesser status would have simply

'disappeared', but instead he was exiled to Gorky with his wife, and the news of his hunger strikes kept leaking out. The nearly terminal sufferings of both Gandhi and Sakharov made the deepest impression on their people and their leaders. Finally, Premier Gorbachev allowed Sakharov to return from Gorky to Moscow, at a critical time when the balance was swung towards democracy rather than towards disastrous international confrontation.

The intellects of both men and their acts of courage were guided by integrity. In Gandhi's opinion,

**As a splendid palace
deserted by its inmates
looks like a ruin,
so does
a man without character,
all his material belongings
notwithstanding.**

The possible consequences of their actions never prevented either man from doing what he perceived to be right. At the very beginning of his involvement with dissident issues, in 1966, Sakharov was asked to sign a statement against Stalin's rehabilitation, a dangerous subject at the time, but he writes simply: "I read his draft letter, found nothing objectionable in it, and added my signature."

However, Gandhi and Sakharov were not helpless martyrs: they both knew how to make the best political use of the sacrifices they made. Another character trait they shared was an openness of manner, combined with a certain detachment, an ability to compromise on the spot. Gandhi was always willing to negotiate:



Gandhi at a spinning wheel during a demonstration in India, June 1925.

**But all my life through,
the very insistence on truth
has taught me to appreciate
the beauty of compromise.**

And they were always willing to try again.

Gandhi said:

**Satisfaction lies in the effort,
not in the attainment.
Full effort
is full victory.**

and

**Life is an aspiration.
Its mission is
to strive after perfection
which is self-realisation.**

Sakharov lived by an epigraph in Goethe's *Faust*:

**He alone is worthy of life and freedom
who each day does battle for them anew.**

We can learn a great deal about a person from the way he judges others. Sakharov provides an insight into his moral convictions when he describes one of his favourite professors as having "absolute intellectual integrity and courage, willingness to re-examine his ideas for the sake of truth, and readiness to take action."

The legacy of the two pioneers may serve to guide humanity in the twenty-first century. Gandhi's warnings and hunger strikes protesting against ethnic cleansing are still important teachings today. Although Sakharov was weakened by his exile and did not live long enough to help guide the new Russia, he left us practical advice in his *Memoirs* and other writings. Thus, the lawyer and the engineer who fought for such similar ideals have left a literature to prepare us for future crises. They may also be role models for us as individuals, setting examples of personal integrity and common sense, and the courage to rise above adversity when it comes. ■

Tom Gehrels is in the Department of Planetary Sciences, University of Arizona, Tucson, Arizona, USA.



Sakharov surrounded by the electorate at the Congress of the USSR People's deputies in May 1989.

Bordeaux Mixture

If at first you can't convince people about the benefits of GM crops — cheat.

Charles Dexter Ward

I'm spraying my tomatoes with bordeaux mixture and it feels great. My wife says I do the tomatoes a disservice, dousing them with Bergerac, when our pension could easily spare Clydebank Cabernet. But the tomatoes love it. No sooner do I get to their row with the sprayer, than their desiccated leaves flush with green; their blooms perk up; their ripening fruits blush with a richer glow. They love me, my tomatoes, and I love them back. Today is special — it's 2090, and my tomatoes and I are celebrating the safe passage of our life-giving Sun through yet another total eclipse.

Anyway, while I'm up there schmoozing, I think how far we've come in just one lifetime. Once upon a time, tomatoes were monotonous plants that took a lot of looking after; constant watering and spraying against greenfly and rot. Today's GM tomatoes are as different from the crop of my youth as the einkorn and emmer they harvested in the Fertile Crescent with obsidian sickles.

I've got rows of them — tomatoes, that is — all the latest kinds: juicy blue ones as big as canteloupes; fluorescent orange fruits the size of pinheads but as hot as habañeros; long thin ones like cucumbers; tetrahedral ones; ones with edible roots; ones that grow like trees which I harvest like apples. And they look after themselves, pretty much; they farm their own mycorrhizae, nurse their own symbionts, kill off the weeds with endogenous antibiotics, and suck what little water they need out of the air. No, I only spray my tomatoes with bordeaux mixture because they enjoy it. They want me to, and, willingly, I oblige.

These days, the word 'tomato' seems almost redundant, as everything else in the garden has undergone much the same kind of transformation. If every plant can be made into anything you want, and made to taste like any other crop, it rather breaks down the barriers. If you have lettuces that look like onions and taste like lemon meringue pie, who cares about horizontal gene transfer?

You'd think that this uncertainty about what's what in the garden would worry me, given that I've always seen myself as something of a tomato connoisseur and never knew my onions. What if I found myself growing an aubergine by mistake, an aubergine that looks like a tomato? But the fact is I don't care: the tomatoes themselves see to that. It gives me such a thrill to see them practically whoop with solanaceous pleasure



as they see me advancing up the garden; such a feeling of contentment as I can hardly describe.

It hardly seems possible that, less than a century ago, people objected so violently to genetic modification, when subsequent history shows it to have been such a wonderful innovation. How silly it all now seems: all those people who trashed test crops seem, in retrospect, like those weavers who broke up to power looms. But then, I was doing some of the modification, so perhaps I'm biased.

Now I'm long since retired, and the company I was working for back in 2007 — when everything changed — has gone the way of Microsoft and Tharsis Telomerase, I can tell all. In 2007, GM was so unpopular with the public that bioscience companies had to fund research and development almost in secret. Progress advanced in giant steps, but all behind the scenes. Unknown to the public, there were plants that did everything except talk back; plants that created their own self-sufficient ecospheres. A few were dropped on the martian South Pole. There were no announcements, no press releases. I hear that a few small stands of martian maize still thrive.

Then a few of us at the lab hit on an idea. We transfected maize with genes for human pheromones. With our corporate heads, we thought that this would do wonders for brand loyalty. The thing is, human pheromones influence behaviour subconsciously. To tell people what we were doing would defeat the object, wouldn't it? Late one night we planted a stand of GM maize in California (I forget exactly where) and within weeks there were activists pounding the door of the Capitol in Sacramento demanding GM crops. Success breeds success — we got the same encouraging results with courgettes in Chihuahua, tomatoes in Thailand and greengages in Glasgow.

But that was long ago, and anyway, when I'm up here with the tomatoes, all that matters is the continuous present, when I am surrounded by the rapturous cacophony of my gorgeous plants — all mine — the plants I love and that love me so much in return, filling the green ether with triumphant shouts of radiant joy.

*Charles Dexter Ward hopes to be the first writer-in-residence on the Space Station. A collection of his fiction, *God Among the Robots and Other Stories*, will be published in 2000 by Unicorn Gardens Press.*

JACEY

From forelimbs to two legs

Mark Collard and Leslie C. Aiello

A reanalysis of the wrist bones of early human fossils provides the first good evidence that humans evolved from ancestors who 'knuckle-walked', as chimps and gorillas do today.

Most people think that palaeoanthropologists spend their lives in exotic places prospecting for fossils. Not so — researchers working on human evolution generally spend most of their time in the laboratory seeking new ways of testing ideas about the existing fossil record. At least as many advances in our understanding of human evolution have come from reanalyses of familiar fossils as from the discovery of new ones. The study described by Richmond and Strait on page 382 of this issue¹ is an excellent case in point.

Richmond and Strait reanalysed several well-known early hominid fossils, including the famous skeleton known as Lucy. Comparing the wrists of these fossils with the wrists of modern humans and several other primate species, Richmond and Strait found that two of our earliest fossil relatives, *Australopithecus afarensis* and *Australopithecus anamensis*, exhibit characteristics of the wrist that are seen today only in the African apes. These features are thought to be associated with knuckle-walking, an unusual mode of quadrupedal locomotion in which the fingers are bent and weight is supported on the backs of the second of the three rows of finger bones. In contrast, the wrists of the other early hominid species in their sample,

Australopithecus africanus and *Paranthropus robustus*, are like those of modern humans in that they lack the putative knuckle-walking characteristics (Fig. 1).

Richmond and Strait's study is significant for several reasons. First, it bears on the long-standing debate over the evolutionary history — the 'phylogenetic' relationships — of modern humans (*Homo*), chimpanzees (*Pan*) and gorillas (*Gorilla*). Genetic analyses overwhelmingly indicate that chimpanzees and humans are more closely related to one another than either is to gorillas². Until the study by Richmond and Strait, however, the anatomical evidence largely ran counter to this conclusion. Many palaeoanthropologists took the knuckle-walking characters of *Pan* and *Gorilla* to be 'shared derived' characters which showed that the two groups of African ape were each other's closest relative³.

The anatomical evidence could be reconciled with the molecular evidence in only two ways. Some researchers thought that the knuckle-walking characters exhibited by *Pan* (Fig. 2, overleaf) and *Gorilla* were independently derived or 'convergent' characters, resulting from natural selection under similar environmental and ecological conditions⁴. This perspective implied that the knuckle-walking characters are of no phylo-

genetic significance. Others argued that the knuckle-walking characters were present in the common ancestor of *Pan*, *Gorilla* and *Homo*, and were simply lost in the lineage leading to *Homo*⁵. According to this view, the knuckle-walking characters of *Pan* and *Gorilla* are 'primitive' features, which, like convergent characters, tell us nothing about phylogeny. Richmond and Strait's results offer strong support for the idea that knuckle-walking characters were present in the common ancestor of modern humans and the African apes, and provide grounds for reconciliation between the molecular and anatomical evidence.

The second point of significance in the new study¹ relates to the phylogenetic relationship between *A. afarensis* and *A. africanus*. For many years, palaeoanthropologists thought that *A. afarensis* was the ancestor of *A. africanus* and the later hominids because its skull has more chimpanzee-like features than that of *A. africanus*. In 1995, however, the discovery at Sterkfontein, South Africa, of a collection of foot bones from a single *A. africanus* individual⁶ complicated the picture. Unexpectedly, the bones indicated that the foot of *A. africanus* was more ape-like than the foot of its putative ancestor *A. afarensis*. This finding was supported by analyses of a newly discovered shin bone from Sterkfontein, which suggested that the lower leg of *A. africanus* was more ape-like than that of *A. afarensis*⁷. It was also supported by an assessment of hominid joint sizes, from which it seemed that *A. africanus* had more ape-like limb proportions than *A. afarensis*⁸.

So it looked like the skulls of these species were telling one story (that *A. afarensis* was the ancestor of *A. africanus* and the later hominids), and their limbs were telling another (that either *A. afarensis* was a side-branch in human evolution and *A. africanus* was the ancestor of the later hominids, or vice versa). The work by Richmond and Strait further complicates the picture: it suggests that *A. afarensis* retained some knuckle-walking features, whereas *A. africanus* did not. It is no longer a case of the skull pointing to one set of phylogenetic relationships, and the postcranial skeleton — everything but the skull — to another. Rather, different parts of the postcranium may not support the same phylogenetic hypothesis.

The third notable aspect of Richmond and Strait's study concerns the interpretation

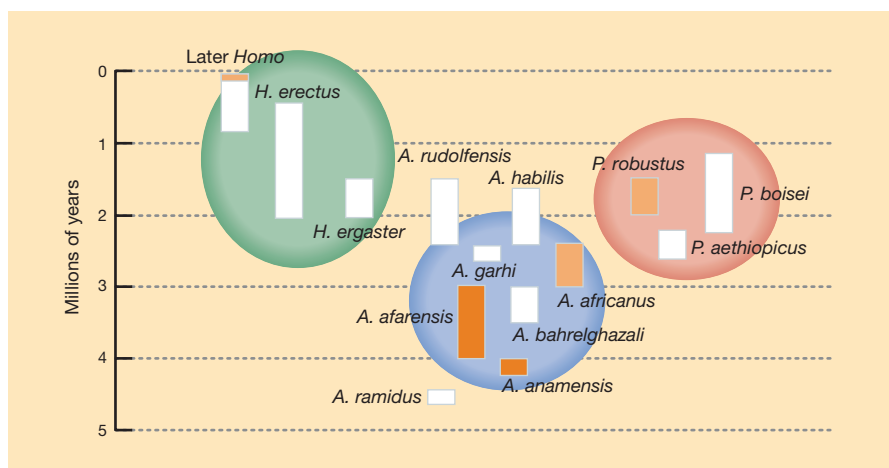


Figure 1 Approximate time ranges of known hominid species. Those species shown in the red circle are robust 'australopithecines' belonging to the genus *Paranthropus* and represent an evolutionary dead-end. Those in the blue circle belong to the genus *Australopithecus*, and those in the green circle are of the genus *Homo*. There is no consensus over the precise phylogeny linking these species. Richmond and Strait¹ have recognized knuckle-walking features in the wrists of those species indicated by dark orange colouring; species indicated in light orange do not have these features. Evidence for the presence or absence of such features in the other species is either non-existent or not yet analysed.



Figure 2 The common ancestor of modern humans and African apes, including gorillas, bonobos and chimps (pictured), may have used a form of locomotion called 'knuckle-walking' when on the ground.

of fossil morphology in terms of function, especially locomotion. Ever since Lucy was discovered in the late 1970s, there has been debate about the locomotor repertoire of *A. afarensis*, the species to which she is assigned.

The basic facts are not in dispute. *A. afarensis* has a combination of traits that is not seen among living primates. In some respects, *A. afarensis* is quite human-like (for instance in the foot structure, non-opposable big toe, and pelvis shape). In others, it is quite ape-like (relatively long and curved fingers, relatively long arms, and funnel-shaped chest).

What can be made of these features? For some researchers, the ape-like characteristics of *A. afarensis* are non-functional retentions from the common ancestor of hominids and the African apes. Here, emphasis is put on the human-like characteristics, and *A. afarensis* is seen as a hominid that walked on two legs and got about in no other way⁹. For others, the ape-like traits are functionally important, and *A. afarensis* is interpreted as using a 'mixed' locomotor repertoire, in which a form of terrestrial bipedalism was combined with an ability to move around effectively in trees¹⁰.

Richmond and Strait add a twist to this debate. They propose that the knuckle-walking features of *A. afarensis* are non-functional retentions from the common ancestor of hominids and African apes. This seems an entirely reasonable position, given that *A. afarensis* shows many traits that are thought to be associated with bipedal locomotion. The alternative idea — that *A. afarensis* combined knuckle-walking, bipedalism and climbing — is somewhat counterintuitive, because it implies the use of two entirely different modes of terrestrial locomotion.

By the same token, however, Richmond and Strait's argument undermines the idea

that *A. afarensis* combined bipedalism with climbing, which many researchers have hitherto considered to be the best interpretation of the evidence. Can we assert that one set of ape-like characters indicates that *A. afarensis* was an able climber, while at the same time arguing that another, equally good, set of ape-like characters is indicative of nothing except the phylogenetic history of *A. afarensis*? If the knuckle-walking characters are considered to be primitive retentions, must not the same hold for the other ape-like characters?

Are our only choices to accept that *A. afarensis* was a striding biped with a large number of non-functional primitive retentions, or to have to take seriously the counterintuitive idea that the locomotor repertoire of *A. afarensis* included forms of bipedalism, climbing and knuckle-walking?

It is difficult to predict how palaeo-anthropologists will react to Richmond and Strait's study. But one thing is certain. It will encourage many researchers to reconsider their assumptions about the phylogenetic and functional implications of bone shape and size in the primates, and most especially in the early hominids.

Mark Collard and Leslie C. Aiello are in the Department of Anthropology, University College London, Gower Street, London WC1E 6BT, UK. e-mails: m.collard@ucl.ac.uk
l.aiello@ucl.ac.uk

1. Richmond, B. G. & Strait, D. S. *Nature* **404**, 382–385 (2000).
2. Ruvolo, M. *Mol. Biol. Evol.* **14**, 248–265 (1997).
3. Andrews, P. & Martin, L. B. *J. Hum. Evol.* **16**, 101–118 (1987).
4. Dainton, M. & Macho, G. A. *J. Hum. Evol.* **36**, 171–195 (1999).
5. Pilbeam, D. R. *Mol. Phyl. Evol.* **5**, 155–168 (1996).
6. Clarke, R. & Tobias, P. V. *Science* **269**, 521–524 (1995).
7. Berger, L. R. & Tobias, P. V. *J. Hum. Evol.* **30**, 343–348 (1996).
8. McHenry, H. M. & Berger, L. R. *J. Hum. Evol.* **35**, 1–22 (1998).
9. Latimer, B. & Lovejoy, C. O. *Am. J. Phys. Anthropol.* **78**, 369–386 (1989).
10. Susman, R. L., Stern, J. T. & Jungers, W. L. *Folia Primatol.* **43**, 113–156 (1984).

Quantum optics

Tricks with a single photon

Peter Zoller

Laser light can trap atoms. But can the light field of a single photon hold an atom in free space? Because it was the development of the laser as an intense light source that made optical trapping and cooling of atoms practical, this seems to be an essentially impossible task. But now a group from the Max-Planck Institute for Quantum Optics in Munich headed by G. Rempe reports an experiment¹ (on page 365 of this issue) in which a single photon stored in a microcavity produces an optical force sufficient to trap an atom. Together with the related experiments on trapping and cooling of single atoms in optical cavities by H. J. Kimble's group at Caltech², these recent advances mark a new generation of cavity quantum electrodynamics (QED) experiments^{1–5}. Such work opens up exciting new

avenues for basic research in quantum physics with single atoms and photons, and suggests fascinating future applications in quantum information processing.

When a laser beam is focused, an atom can experience an optical force so that it is attracted to the centre of the light focus. The same concept underlies 'optical tweezers', which are used to manipulate much more macroscopic particles than atoms with light fields. For intense laser beams, the optical force can be strong enough to overcome gravity. The stronger the light beam, the deeper is the optical potential that holds the atom. In addition, for an atom to remain bound in the laser focus, its thermal energy must be sufficiently small that it cannot escape. In fact, the atoms must be extremely cold in comparison with gases at room

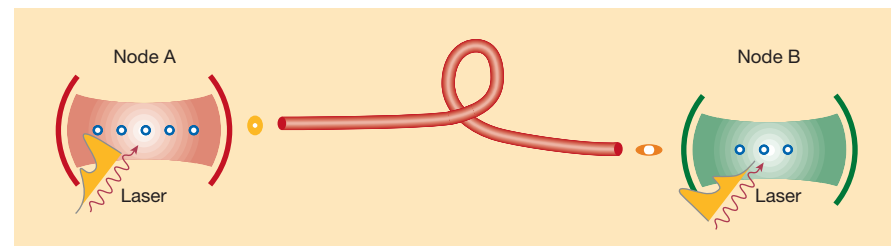


Figure 1 A quantum network based on cavity quantum electrodynamics (QED). Quantum communication channels connect the spatially separated nodes. Each node is a quantum processor that stores and processes quantum information locally. Atoms in the cavity provide the quantum memory, laser pulses play the role of quantum gates by exchanging cavity photons, and the cavities themselves become optical interconnects with a fibre linking the cavities. Exchange of information between the nodes of the network is accomplished by way of quantum channels. New techniques to confine atoms within optical cavities^{1–3} bring such quantum networks closer to reality.

temperature, and it was only the development of laser cooling techniques during the past two decades, allowing atoms to be cooled to microkelvin temperatures, that made optical trapping of atoms possible. These achievements were honoured with the award of the 1997 Nobel Prize in Physics to Chu, Cohen-Tannoudji and Phillips.

Quantum physics teaches us that the light field is quantized; that is, light comes in energy units of $E = h\nu$, where h is Planck's constant and ν is the light frequency. These light quanta are called photons. The granular quantum structure of light becomes visible only when we deal with extremely weak light fields consisting of a few photons. In this sense, a single photon represents the weakest possible light field. So how can the light intensity of a single photon be sufficient to trap an atom? The key idea is to store the photon in an optical cavity consisting of two high-quality mirrors. The smaller the volume in which we confine the photon, the larger is the energy per volume, and it is the corresponding strong electric field that is seen by the atom and determines the optical force. Trapping an atom with a single photon can thus be achieved using microcavities with extremely high-quality optical mirrors, where the distance between the mirrors is only ten or twenty times the optical wavelength. This is the basis of the Max-Planck and Caltech experiments¹⁻³.

How are atoms actually captured inside the cavity, and how do we know that it is there? The cavity mirrors are not perfect, and a single photon will be lost from the cavity after a short time. These photons must therefore be continuously replenished by a pumping laser that puts the photon back into the cavity. But a unique feature of such cavity QED experiments is that the light field leaking out of the cavity allows us to observe the motion of the atom in real time inside the cavity. So we can 'see' an atom entering the cavity, and this enables a feedback loop to change the parameters guiding the laser, which can then capture the atoms.

The latest developments in cavity QED with optical and microwave photons should be seen in a much broader context. An important new direction in modern physics has been the attempt to control processes in individual quantum systems on a quantum by quantum level. One of the long-term hopes of these endeavours is to develop the ability to assemble quantum devices as composite controllable quantum systems. Quantum optics has taken a leading role in this area, and cavity QED and laser-cooled trapped ions are its prime examples. The recent achievements in cavity QED¹⁻⁵ and ion traps⁶ now set the stage for future developments, ranging from basic studies of quantum physics all the way to possible applications.

Among the most exciting prospects are

the possibilities offered by cavity QED to build quantum information processors and to connect these devices in an optical quantum network⁷. Quantum computers require the storage of information in a set of two-level systems (called quantum bits or qubits), the processing of this information using quantum gates, and a final readout by a quantum-state measurement.

In a typical cavity QED set up, long-lived internal states of atoms can play the role of reliable quantum memory for storing the qubits. Quantum operations on a single qubit can be performed by coupling two atomic states with laser pulses. Joint operations on two qubits require the controlled exchange of a cavity photon between two atoms, where the cavity mode plays the role of a quantum data bus. This mechanism for a two-bit gate relies directly on the strong coupling of single photons with single atoms.

Although isolated atoms represent an ideal quantum memory, optical photons propagating in fibres are a fast and reliable

way to transmit quantum information to distant nodes of a network (Fig. 1). Cavity QED provides a natural setting for the interface between atoms and photons in the system in the form of optical interconnects. The primary technical obstacle to achieving these goals is the need to develop techniques to confine atoms within cavities. The recent experiments by the Caltech and Max-Planck groups are important milestones in this direction. ■

Peter Zoller is at the Institute for Theoretical Physics, University of Innsbruck, 6020 Innsbruck, Austria.

e-mail: peter.zoller@uibk.ac.at

1. Pinkse, P. W. H., Fischer, T., Maunz, P. & Rempe, G. *Nature* **404**, 365–368 (2000).
2. Hood, C. J., Lynn, T. W., Doherty, A. C., Parkins, A. S. & Kimble, H. J. *Science* **287**, 1447–1453 (2000).
3. Ye, J., Vernooij, D. W. & Kimble, H. J. *Phys. Rev. Lett.* **83**, 4987–4990 (1999).
4. Nogues, G. *et al.* *Nature* **400**, 239–242 (1999).
5. Varcoe, B. T. H., Brattke, S., Wedinger, M. & Walther, H. *Nature* **402**, 743–746 (2000).
6. Sackett, C. *et al.* *Nature* **404**, 256–259 (2000).
7. Cirac, J. I., Zoller, P., Kimble, H. J. & Mabuchi, H. *Phys. Rev. Lett.* **78**, 3221–3224 (1997).

Neurodegenerative diseases

Parkinson's pathology in a fly

Christian Haass and Philipp J. Kahle

Human life expectancy has increased dramatically over the past century. But can we really enjoy our extra years? Diseases such as Alzheimer's and Parkinson's, characterized by the death of certain nerve-cell populations and consequent dementia or movement defects, threaten more and more elderly people in the Western world. Many of the molecules behind these diseases have been identified, but we still lack a coherent understanding of the processes involved. On page 394 of this issue¹, however, Feany and Bender offer a new way to tackle the problem. They have developed a model of Parkinson's disease in the fruitfly *Drosophila melanogaster* that reproduces many of the features of the human disorder. The hope is that, with this model, we can learn much about the disease and, perhaps, new ways to treat it.

Parkinson's disease is the second most common neurodegenerative disorder. The pathological hallmark of the disease is the accumulation of fibrous protein deposits in neuronal cytoplasm (Lewy bodies) and nerve fibres (Lewy neurites) in the brain. These deposits may interfere with normal neuronal function. Selective death of the neurons that normally secrete the neurotransmitter dopamine results in a movement disorder that is characterized by muscle rigidity and resting tremor. The neurons that are affected in patients with Parkinson's disease are found in the regions of the brain

called the substantia nigra and the locus coeruleus (and form the nigrostriatal dopamine system).

Mice, flies and worms that have been engineered to express human genes linked to Alzheimer's disease have already yielded invaluable insight into the molecular mechanisms underlying this disorder², and have even offered exciting possibilities for therapy³. Until recently, it was difficult to take the same tack for Parkinson's disease, as genes linked to its onset or progression were unknown. Three years ago, however, by studying an Italian-Greek family in which Parkinson's disease was inherited, Polymeropoulos *et al.*⁴ identified a critical mutation in the gene encoding α -synuclein. This mutation resulted in alanine being replaced by threonine at amino-acid position 53 of the protein (A53T mutation). A second mutation, resulting in the substitution of alanine with phenylalanine at position 30 (A30P mutation), was later identified in a German family by Krüger *et al.*⁵ (Fig. 1a, overleaf).

Interestingly, α -synuclein, a phosphorylated protein⁶ normally found in presynaptic neurons⁷, is the main component of Lewy bodies⁸ (Fig. 1b, overleaf). Ablation of the gene encoding α -synuclein in mice results in functional deficits of the nigrostriatal dopamine system⁹, providing further evidence for a connection of α -synuclein with Parkinson's disease. In familial Alzheimer's disease, mutations in several distinct genes

accelerate the aggregation and deposition of the amyloid- β peptide¹⁰, the pathological hallmark of this disease. Likewise, it was expected that the α -synuclein mutations might enhance the formation of Lewy bodies, and indeed, early *in vitro* experiments showed that α -synuclein could self-aggregate and form fibres^{11,12}.

These findings provided the basis for the generation of an animal model that recapitulates the hallmarks of human Parkinson's disease, a task now achieved using *Drosophila* by Feany and Bender¹. The authors expressed the gene encoding human α -synuclein (as a 'transgene') in all *Drosophila* nerve cells. This led to an age-dependent loss of dopamine-secreting (dopaminergic) neurons; other neurons appeared to be largely unaffected. Death of dopaminergic cells was observed upon expression of wild-type α -synuclein or of either of the two mutants (A53T or A30P) identified previously. Strikingly, Feany and Bender also found that some neurons accumulated intracellular aggregates that closely resembled the Lewy bodies found in Parkinson's disease patients. Most important, the inclusions were composed of α -synuclein filaments 7 to 10 nanometres in diameter — again, characteristics of human Lewy bodies.

So, the transgenic flies showed many of the physical traits of patients with Parkinson's disease. But did the flies also show the

behavioural characteristics seen in humans? To address this question, Feany and Bender used a climbing test — an assay originally designed to follow age-related changes in *Drosophila*. When tapped down to the bottom of a vial, flies rapidly climb up to the top in what is known as a negative geotactic response. Feany and Bender found that young flies overexpressing α -synuclein performed well in this test, but aged transgenic flies frequently fell back to the bottom. Flies expressing wild-type α -synuclein or the A53T mutant performed similarly to each other in this test, but locomotor defects were more severe in flies expressing the A30P mutant. This may indicate that the A30P mutant is more aggressive than the A53T mutant, although the possibility that there are small differences in transgene expression cannot be excluded. Nevertheless, this fly model clearly replicates many of the characteristic features of Parkinson's disease.

Many laboratories are also trying to generate a transgenic mouse model of Parkinson's disease. Masliah *et al.*¹³ have now developed transgenic mice that express wild-type α -synuclein under the control of the promoter of the platelet-derived growth factor gene, which is expressed in all neurons. These mice accumulate inclusions composed of α -synuclein, although the inclusions do not contain fibres. The amorphous inclusions are seen not only in neuronal cell

bodies but also in nuclei of transgenic mice, where they do not occur in Parkinson's disease. Some neurons appear to react with antibodies directed towards a protein called ubiquitin — a feature of many classical Lewy bodies in the brains of Parkinson's patients. The transgenic animals also show dopaminergic deficits. Amounts of tyrosine hydroxylase (an enzyme required for the biosynthesis of dopamine) are significantly reduced in animals with excessive α -synuclein expression. Moreover, behavioural tests revealed locomotor impairments in these mice.

In other mouse models, pan-neuronal expression of the A30P α -synuclein mutant led to pathological defects in neuronal extensions (neurites)^{14,15}. Classical Lewy bodies have not yet been observed in these models. But studies of these mice have revealed something interesting about the A30P mutant. Wild-type α -synuclein has been reported to bind to cell membranes *in vitro*, whereas mutant α -synuclein does not¹⁶. In these mouse models, however, the A30P mutant does not seem to lose this ability entirely, so it shows a toxic gain, rather than loss, of function¹⁵.

Now that models for Parkinson's disease have been developed in vertebrate and invertebrate animals, a broad range of investigative options has been opened up. Researchers will be able to study the influence of pathological protein expression on neuronal function and animal behaviour. A fundamental question can be addressed: whether or not deposition of α -synuclein causes selective neuronal death in Parkinson's disease. And finally, it will now be possible to screen for compounds that might interfere with pathological events in Parkinson's disease. In this respect, the short generation time of flies makes them invaluable tools for drug screening, and our in-depth knowledge of their genetics will aid in the identification of genes that affect the onset and progression of this debilitating disease. ■

Christian Haass and Philipp J. Kahle are at the Adolf-Butenandt-Institute, Department of Biochemistry, Ludwig-Maximilians-University Munich, Schillerstrasse 44, 80336 Munich, Germany. e-mails: chaass@pbm.med.uni-muenchen.de pkahle@pbm.med.uni-muenchen.de

1. Feany, M. B. & Bender, W. W. *Nature* **404**, 394–398 (2000).
2. Price, D. L., Sisodia, S. S. & Borchelt, D. R. *Science* **282**, 1079–1083 (1998).
3. Schenk, D. *et al.* *Nature* **400**, 173–177 (1999).
4. Polymeropoulos, M. H. *et al.* *Science* **276**, 2045–2047 (1997).
5. Krüger, R. *et al.* *Nature Genet.* **18**, 106–108 (1998).
6. Okochi, M. *et al.* *J. Biol. Chem.* **275**, 390–397 (2000).
7. Maroteaux, L. & Scheller, R. H. *Brain Res. Mol. Brain Res.* **11**, 335–343 (1991).
8. Spillantini, M. G. *et al.* *Nature* **388**, 839–840 (1997).
9. Abeliovich, A. *et al.* *Neuron* **25**, 239–252 (2000).
10. Selkoe, D. J. *Nature* **399**, A23–A31 (1999).
11. Conway, K. A., Harper, J. D. & Lansbury, P. T. *Nature Med.* **4**, 1318–1320 (1998).
12. Narhi, L. *et al.* *J. Biol. Chem.* **274**, 9843–9846 (1999).
13. Masliah, E. *et al.* *Science* **287**, 1265–1269 (2000).
14. Hsiao, K. *Ann. NY Acad. Sci.* (in the press).
15. Kahle, P. J. *et al.* *Ann. NY Acad. Sci.* (in the press).
16. Jensen, P. H. *et al.* *M. J. Biol. Chem.* **273**, 26292–26294 (1998).
17. Ueda, K. *et al.* *Proc. Natl Acad. Sci. USA* **90**, 11282–11286 (1993).

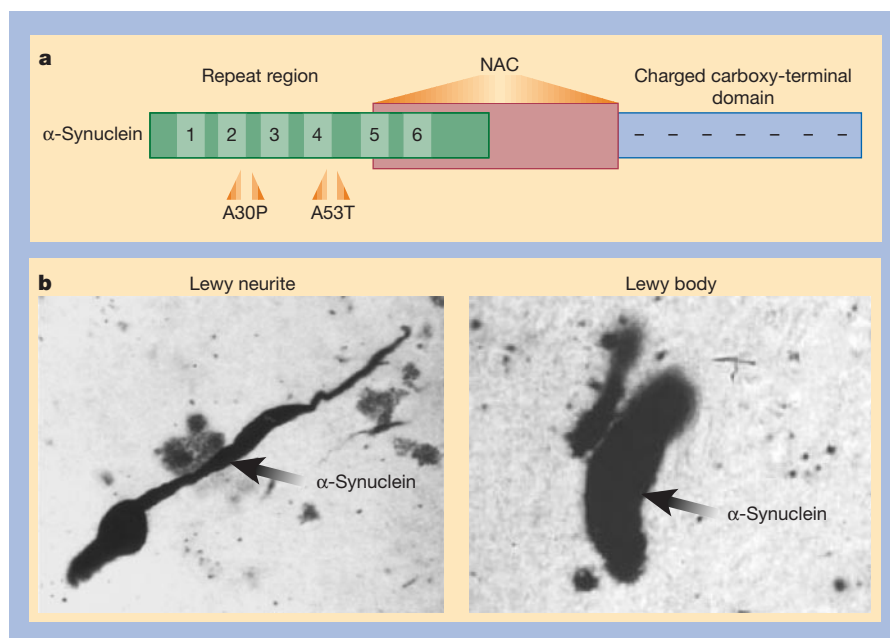


Figure 1 A central role for α -synuclein in Parkinson's disease. **a**, The structure of the α -synuclein protein. The potentially lipid-binding amino-terminal domain consists of an imperfectly repeated motif, represented by the numbers 1–6. The NAC domain corresponds to the non-amyloid component originally isolated from senile plaques of patients with Alzheimer's disease¹⁷. The two mutations associated with familial Parkinson's disease are indicated. In the A30P mutant, alanine at amino-acid position 30 is mutated to proline; in the A53T mutant, alanine at position 53 is replaced by threonine. **b**, Lewy neurites (left) and Lewy bodies (right) composed of α -synuclein are the characteristic lesions observed in the brains of patients with Parkinson's disease. Similar inclusions of α -synuclein are observed in Feany and Bender's fly model of Parkinson's disease¹. These sections were stained with a polyclonal antibody to human α -synuclein⁶.

A needle in a cosmic haystack

Isabelle A. Grenier

When the needle is a powerful source of gamma rays and the haystack is a region of the sky filled with thousands of stars, nebulae and galaxies then, as traditional wisdom dictates, a lot of time and effort is required to find the needle. Over the past few years the EGRET telescope, on board the Compton Gamma-Ray Observatory, has revealed a surprisingly large number (about 170) of gamma-ray sources above 100 MeV (megaelectron volts), most of which apparently belong to our Galaxy¹. Gamma rays are the most energetic form of electromagnetic radiation. They are usually absorbed in the Earth's upper atmosphere (only the most energetic can be detected by ground-based telescopes), so nearly all gamma-ray observations have to be carried out in space.

A handful of gamma-ray sources were found in the 1970s and 1980s by the SAS-2 and COS-B satellites, but however bright and stable they appear, they remain an enigma. The difficulty with identifying the objects behind these extraordinary sources results from the poor positional accuracy of past and present high-energy telescopes, which is of the order of a degree, or twice as large as the apparent size of the Sun. No clear picture has yet emerged as to the nature or even the variety of objects involved: they could be neutron stars, black holes, supernova remnants or massive stars. On page 363 of this issue, Gehrels *et al.*² present important evidence for two distinct

populations among the persistent unidentified sources in the Galaxy: bright sources spread within the Galactic disk and a score of fainter objects that appear to lie in the solar neighbourhood.

Along the Milky Way, most of the unidentified sources are clearly seen in the direction of active star-forming sites^{1,3,4}, which shelter many possible gamma-ray emitters, such as pulsars (isolated or in binary systems), powerful shock waves from supernova remnants, massive stars with supersonic winds, and even black holes. Source identification in these environments is frustratingly hindered by an overabundance of candidates. Current explanations⁵ mostly focus on pulsar activities, either on radiation beams from the neutron star or on emission from the relativistic wind and synchrotron nebula it powers. But higher-precision observations with the Gamma-ray Large Area Space Telescope (GLAST), due to be launched in 2005, might yet surprise us.

Away from the Galactic plane, the opposite situation prevails. There are no conspicuous counterparts at other wavelengths near many of the gamma-ray sources. A famous example is that of Geminga, a source nicknamed "it is not there" in Milanese dialect, finally identified after 20 years as the closest known pulsar. Moreover, it is not easy to use the collective properties of the sources to get clues to their nature because of strong observational biases. Modelling the latter means guessing at the identity of the

parent population. So it is important that Gehrels *et al.*² have shown, without modelling, that the stable sources away from the Galactic plane are not artefacts and are distinct in their intensity and spectral properties from the sources in the plane. Earlier studies⁵ showed that their distribution in the sky is significantly correlated with the nearby star-forming Gould belt and with no other obvious Galactic structure. Together, these findings provide compelling evidence that the belt is an important birthplace of gamma-ray sources.

The Gould belt⁶ — a large expanding disk or ring of young stars and gas a thousand light years in radius — is unexpectedly inclined at roughly 20° to the Galactic plane (Fig. 1). The Sun lies halfway to the rim. This structure was created by a violent event of unknown origin, possibly a series of supernova explosions or the impact of a high-velocity cloud onto the Galaxy 30 to 40 million years ago. It is swarming with young stars formed because of the belt's expansion into the surrounding medium. One idea is to explore the possibility⁷ that gamma-ray sources in the belt, in particular those singled out by Gehrels *et al.*, are pulsars born there over the past few million years. The numerous massive stars formed in the starburst region have indeed had time to evolve, explode and produce neutron stars or black holes.

The current stellar population of the belt indicates that it has recently produced many times more supernovae locally than throughout the Galaxy⁷. There is no evidence yet of gamma radiation above 100 MeV from a stellar black hole, but there is from eight pulsars. Seven of them are bright, relatively young (less than 3×10^5 years) objects seen at distances of a few thousand light years, well within the Galaxy. The other one, Geminga, is the first example of a gamma-ray pulsar found in the belt. Why Geminga emits most of its energy as high-energy gamma-rays is not understood⁸.

The Gould belt sources might constitute a population of older, closer and fainter pulsars that would considerably broaden our perspective of particle acceleration and radiation processes inside pulsar magnetospheres. Were it not for its proximity, the extreme faintness of Geminga at wavelengths other than gamma rays would have prevented its pulsation from being detected. But the next-generation telescope, GLAST, will allow direct searches for periodic signals in the gamma-ray range.

As possible relics of supernova explosions in the Gould belt, the gamma-ray sources highlight how dynamic the cycle of matter through star formation and death is in our neighbourhood — a lively drama that eludes us because of the stillness of the optical sky to our naked eyes. The high-energy sky will undoubtedly provide more

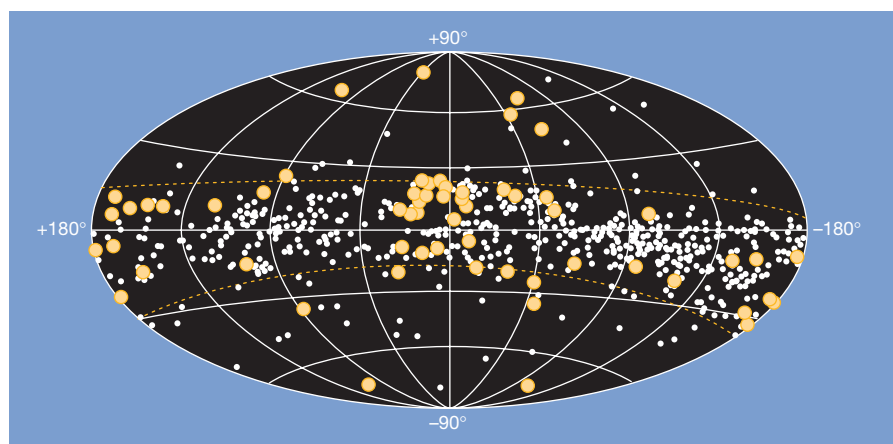


Figure 1 Unidentified gamma-ray sources in the Gould belt. A lane of massive stars (white dots), most of which are less than 20 or 30 million years of age, arches across the sky. This starburst disk, named after Benjamin A. Gould, dominates our cosmic neighbourhood up to a thousand light years away, and includes many of the brightest, most conspicuous stars in the sky. Most of the stable, unidentified gamma-ray sources found by the Energetic Gamma Ray Experiment Telescope (EGRET) at mid-latitudes (yellow dots) closely follow the curve and apparent width of the belt. Gehrels *et al.*² show that these mid-latitude sources are distinct from the population of unidentified sources in the Galactic plane, and are likely to be associated with the Gould belt.

surprises, such as the existence of a third population of Galactic gamma-ray sources, which was announced at a symposium last August⁵. The new X-ray satellites, Newton (XMM) and Chandra, the future gamma-ray missions, Integral and GLAST, and new ground-based telescopes under construction will soon move us from an era of early exploration to one of understanding the extreme environments of neutron stars and black holes in our neighbourhood. ■

Isabelle A. Grenier is at the Université Paris VII and Service d'Astrophysique CEA-Saclay, Centre d'Etudes de Saclay, 91191 Gif/Yvette, France.

e-mail: isabelle.grenier@cea.fr

1. Mukherjee, R., Grenier, I. A. & Thompson, D. J. in *Proc. 4th Compton Symp.* (eds Dermer, C. D., Strickman, M. S. & Kurfess, J. D.) 394–406 (American Institute of Physics, New York, 1997).
2. Gehrels, N., Macomb, D. J., Bertsch, D. L., Thompson, D. J. & Hartman, R. C. *Nature* **404**, 363–365 (2000).
3. Yadigaroglu, I. A. & Romani, R. *Astrophys. J.* **476**, 347–356 (1997).
4. Romero, G. E., Benaglia, P. & Torres, D. F. *Astron. Astrophys.* **348**, 868–876 (1999).
5. Grenier, I. A. in *Proc. 6th Towards a Major Atmospheric Cherenkov Detector Symp.* (eds Dingus, B. L., Salamon, M. H. & Kieda, D. B.) (American Institute of Physics, New York, in the press).
6. Pöppel, W. G. L. *Fund. Cosm. Phys.* **18**, 1–271 (1997).
7. Grenier, I. A. & Perrot, C. in *Proc. 26th Int. Cosmic Ray Conf.* **3**, 476–479 (1999).
8. Thompson, D. J., Harding, A. K., Hermsen, W. & Ulmer, P. in *Proc. 4th Compton Symp.* (eds Dermer, C. D., Strickman, M. S. & Kurfess, J. D.) 39–56 (American Institute of Physics, New York, 1997).

X-ray optics

Making hard light sharper

Roland Smith

Light provides us with an incredible range of tools for probing and manipulating nature. We use it to catch and cool atoms to near absolute zero¹, and to heat matter to the extreme conditions found in supernovae². The electromagnetic spectrum (light in the most general sense) extends from longer radio wavelengths through the infrared and visible, and into the shorter ultraviolet, X-ray and gamma-ray range. Over some of this range we have ways of

controlling and harnessing light by using lenses to focus it, mirrors to change its direction and prisms or gratings to split it into its constituent wavelengths. But the tools that we take for granted in the visible region of the spectrum (where the wavelength of light is about 5×10^{-7} m) become progressively harder to build as we move to shorter and shorter wavelengths.

In the 'hard' X-ray range, where the wavelength of light might be 5,000 times smaller

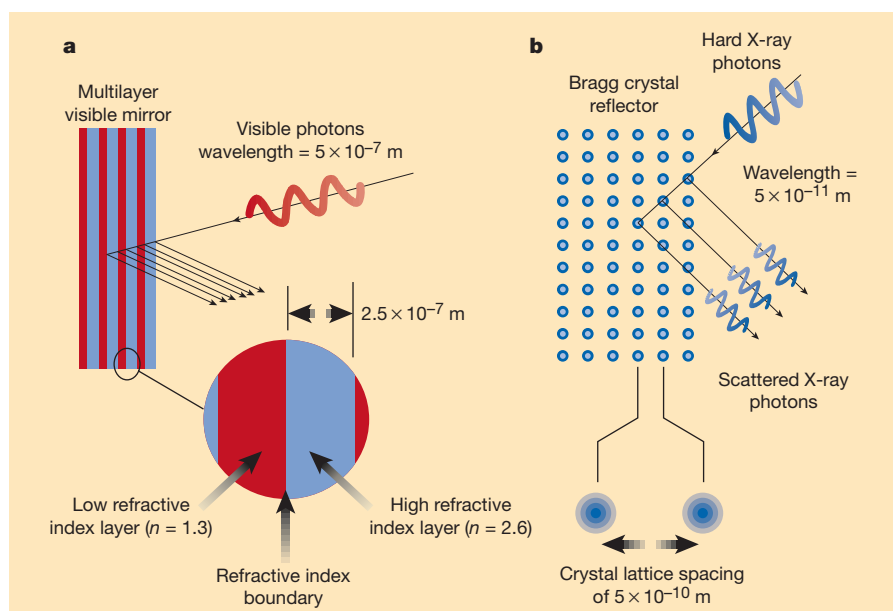


Figure 1 Visible and X-ray optics. **a**, For visible wavelengths, high-quality optical mirrors and filters are often constructed from stacks of material such as cryolite (Na_3AlF_6) and titanium dioxide (TiO_2) with very different refractive indices. Stacks are made with layers half a wavelength thick for mirrors and a quarter of a wavelength thick for filters and anti-reflection coatings. Interference between light reflected at the various layer boundaries gives increased reflection or transmission for a specific wavelength and input angle. **b**, At 'hard' X-ray wavelengths ($< 10^{-10}$ m), all materials have a refractive index of nearly one so multilayer optics of this type cannot be constructed. But Bragg scattering from the regular arrays of atoms in a crystal lattice (where the interatomic spacing might be $\sim 5 \times 10^{-10}$ m) can be used to select and reflect X-rays in a similar way.



100 YEARS AGO

I have nowhere seen an account of a very remarkable display of these meteors visible here (in Shanghai) on the morning of November 15, 1886. Though the date is distant, it may be of use to record it, as it may throw light on the conditions of the orbit. I was sleeping in a room with an almost due north exposure looking into an open compound, and chanced to wake up about three in the morning, when I saw a number of meteors flashing across the window. I got up on recollecting the date, and for about an hour witnessed the most brilliant pyrotechnic display I have ever seen. The meteors were flying in every direction from the radiant point in numbers past all calculation, and the intensity of the shower was kept up without intermission the whole of the time I was gazing. I expected to hear from other quarters an account of the phenomenon, and was much surprised to find it had apparently not been noticed elsewhere... As much stress is laid on the appearance of the meteors in Europe in 1833 and 1866, the shower may be some of interest. From *Nature* 22 March 1900.

50 YEARS AGO

The council of the U.S. National Academy of Science has submitted to President Truman a statement relating to certain provisions of the National Science Foundation Bill HR 4846 as passed by the House of Representatives on March 1... The statement continues: "One set of protective measures is rightly aimed at the security of information vital to the national defense. Scientific and technical information has come to play an important role in this defense. We are gravely disturbed, however, to see that security measures are being extended widely over the scientific life of the country, even in those areas remote from possible military application. This development will defeat the growth of science by inhibiting the free exchange of information so vital to it, by discouraging the bravest and most original minds, and by the pervasive threat of irreparable injury to individuals inherent in all counter-intelligence measures..." The Academy is particularly concerned about the amendment providing "for FBI investigation and clearance of every person who is to be awarded a scholarship or fellowship under the terms of the bill"... To submit large numbers of young persons to such investigations is unnecessary. From *Nature* 25 March 1950.

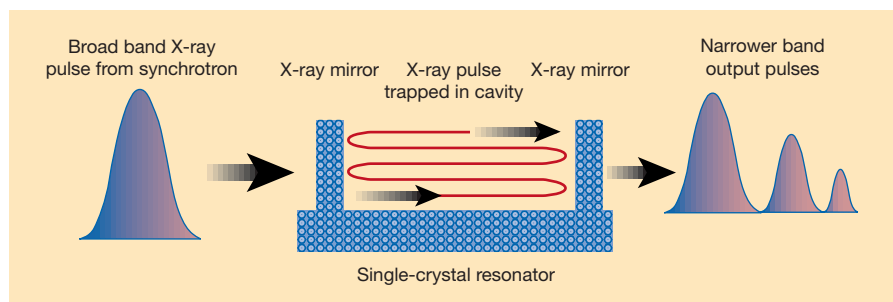


Figure 2 In the device created by Liss *et al.*³, a short-duration (1×10^{-10} second) 'hard' X-ray pulse from a synchrotron is injected into a cavity formed from a single silicon crystal. Some of the pulse becomes trapped in the cavity and makes multiple round trips (14 successive bounces are seen in the experiment). As it does so, the spread of wavelengths in the pulse becomes narrower and narrower because the Bragg condition for highest reflectivity is satisfied for only a small range of wavelengths.

than for visible light, the construction of something as simple as a mirror requires a technical *tour de force*. On page 371 of this issue, Liss *et al.*³ describe a device rather more ambitious than a single mirror. They have built a cavity for hard X-rays from two thin plates of crystalline silicon, and used this to trap and control short pulses of X-rays from a synchrotron. This may well lead to the development of very narrow band filters for hard X-rays.

The X-rays Liss *et al.*³ work with have a wavelength of 7.8×10^{-11} m and are produced in short (1×10^{-10} second) bursts from an electron storage ring at the European Synchrotron Radiation Facility in Grenoble, France. To understand what makes their ability to trap these X-rays so impressive, we have to appreciate why it is difficult to work in this extreme part of the spectrum. In general, the refractive index of a medium, n , describes how much slower light travels in a material than it does in a vacuum, and values of n from 1.3 to 2.6 are common for visible light passing through materials such as glass. When light passes through a boundary between regions of different refractive index, it can be reflected or its path can be changed. This simple fact underlies the operation of lenses, multi-layer coatings and optical fibres used to control visible light, and explains everyday phenomena such as why you can see reflections in a pane of glass.

Unfortunately, in the X-ray region of the spectrum all materials have refractive indices very close to one, so these important boundary effects simply disappear. A second problem is that many useful optical devices (for example, narrow-band filters and high-power laser mirrors) use several layers of materials of different refractive index, with each layer having a very precise thickness of one quarter or one half of the wavelength of light (Fig. 1a). For visible light, constructing such devices is technically feasible, but in the hard X-ray range, building such structures is extremely difficult.

At this point, nature comes to our rescue in the form of neat, regularly spaced arrays of

atoms conveniently packaged as crystals. Individual X-rays can bounce off the atoms in a crystal and the coherent addition of scattered photons from a regular lattice can result in reflection in a preferred direction through a process known as Bragg scattering (Fig. 1b). This process is often used to unravel the detailed structure of matter by recording an X-ray diffraction pattern from a crystalline sample. But it can also be used to select and reflect specific X-ray wavelengths. The device created by Liss *et al.*³ involves Bragg scattering between two wafers of silicon, which were machined from a single block of crystal to keep them precisely aligned. A pulse of X-rays from the synchrotron is loaded through the back of one of the slightly leaky mirrors (which are 70% reflective), and makes multiple bounces inside the cavity (Fig. 2).

Optical cavities form the basis of important devices such as the laser. But it seems unlikely that cavities of the type described here will be used in lasers in the foreseeable future owing to the lack of another crucial component: a medium for amplifying light

that can operate at such short wavelengths. X-ray lasers are under continuing development, and work on 'soft' X-ray cavities (with wavelengths a hundred times longer than reported here) has been quite successful⁴. But there is another powerful motivation for building the type of cavity chosen by Liss and colleagues. As light bounces repeatedly between two closely spaced mirrors, specific wavelengths can be selected by changing the mirror spacing. This forms the basis of a device known as a Fabry–Perot etalon⁵, which can be used for extremely precise wavelength control and measurement.

Synchrotrons are extraordinarily useful research tools, but unlike lasers they generate a broad swathe of wavelengths that overlap in space and time. Many techniques that use light from synchrotrons require a device such as a monochromator to select a small band of these wavelengths to work with. Using a hard X-ray Fabry–Perot etalon may help to reduce the width of this band still further. Liss *et al.*³ now plan to construct such a device and hope to use it to narrow the X-ray line width produced by a synchrotron by a factor of ten or more. This would then allow measurements that depend on the sharpness of this X-ray band, for example X-ray protein crystallography and extended X-ray absorption fine-structure spectroscopy, to be made far more precisely. ■

Roland Smith is in the Department of Physics, Imperial College, Prince Consort Road, London SW7 1BQ, UK.

e-mail: r.a.smith@ic.ac.uk

1. Cohen-Tannoudji, C. N. & Phillips, W. D. *Phys. Today* **43**, 33–40 (1990).
2. Remington, B. A., Arnett, D., Drak, R. P. & Takabe, H. *Science* **284**, 1488–1493 (1999).
3. Liss, K. D. *et al. Nature* **404**, 371–373 (2000).
4. Ceglio, N. M., Stearns, D. G., Gaines, D. P., Hawryluk, A. M. & Trebes, J. E. *Opt. Lett.* **13**, 108–110 (1988).
5. Saleh, B. E. A. & Teich, M. C. *Fundamentals of Photonics* (Wiley, New York, 1994).

Membrane fusion

Changing partners

Chavala M. Carr and Peter J. Novick

Fusion of the membranes that surround cells and their intracellular organelles occurs during a variety of cellular processes. During neurotransmission, for example, neurons release chemical neurotransmitters from intracellular, membrane-bounded containers (vesicles) into the extracellular space, the synapse. This process involves the merging of the vesicle membrane and the presynaptic plasma membrane, allowing the neurotransmitter to be released. Yet before one membrane can merge with another, a series of protein interactions is required in a process that resembles changing partners at a square dance. On page 355 of this issue¹, Misura

and colleagues offer new insight into the relationship between two of these partners.

Central to the mechanism of membrane fusion are the SNAREs, which are proteins, found on the target membrane and on the vesicle, that catalyse membrane fusion by assembling into a complex that pins the two fusing membranes together. A key step in the assembly of the SNARE complex is the activation of syntaxin, a name that encompasses a family of SNARE proteins found on target membranes (the presynaptic plasma membrane in the case of neurotransmission). Thanks to the resolution of the three-dimensional structure of the core of a SNARE complex², our understanding of

membrane fusion is advancing apace to an atomic level; however, the mechanism by which syntaxin is activated remains poorly understood.

Misura and colleagues¹ contribute to this debate by describing the structure of what may be a critical intermediate in the assembly of neuronal SNARE complexes — the complex of syntaxin 1a with another one of its partners, neuronal Sec1 (nSec1). Although Sec1 is essential for membrane fusion *in vivo*, its exact function is not yet clear. By following syntaxin as it steps through the cycle of SNARE-complex assembly and disassembly (Fig. 1), we can get some idea of where Sec1 may enter the dance.

The mechanism of syntaxin activation centres on the accessibility of its so-called H3 helix, a structural feature that must be exposed to allow syntaxin to assemble into SNARE complexes. Accessibility of the H3 helix is regulated by a three-helix bundle structure³ (the inhibitory domain) at syntaxin's amino terminus. When the H3 helix is bound to the inhibitory domain in a putative four-helix bundle structure, SNARE-complex assembly is blocked^{4,5}. In this conformation, syntaxin is 'closed'. Syntaxin is activated when the inhibitory domain is displaced from the H3 helix (step 1 in Fig. 1), resulting in an 'open' conformation to which other SNARE proteins can bind.

Open syntaxin then assembles with other SNAREs into a 'trans'-SNARE complex that holds the two fusing membranes together (step 2). Most prominent among the multiple partners of syntaxin 1a are VAMP, a SNARE on the vesicle membrane, and SNAP-25,

another target-membrane SNARE⁶. The helices from these proteins assemble with the syntaxin H3 helix into a four-helix bundle, forming the *trans*-SNARE complex required for membrane fusion (step 3).

Assembly of *trans*-SNARE complexes is sufficient for the fusion of synthetic membranes⁷, but membrane fusion in the cell requires a number of other conserved proteins, including Sec1. The role of Sec1, which binds to syntaxin⁸, might be to activate syntaxin, so regulating the assembly of SNARE complexes. The structure determined by Misura *et al.*¹ shows nSec1 bound to syntaxin 1a, with syntaxin's H3 helix distorted. In this structure, syntaxin 1a may have been shifted into an intermediate conformation, somewhere between closed and open. But, if nSec1 does indeed activate syntaxin 1a, other components must be required too, because nSec1 alone binds tightly to syntaxin and inhibits SNARE-complex assembly⁹. For this reason, Misura and colleagues speculate that SNARE-complex assembly is stimulated by other factors, which bind nSec1; nSec1 then changes conformation and releases syntaxin in an open conformation. A variety of Sec1-binding proteins have been identified and proposed to activate SNARE-complex assembly in this way. It will be important to see whether, as predicted by this model, SNARE-complex levels are altered in mutants in which Sec1 is non-functional.

Surprisingly, studies of yeast mutants place the essential function of Sec1 after SNARE-complex assembly (step 2) and before membrane fusion (step 3). But the

situation in yeast and neurons may be different: unlike its neuronal counterpart, Sec1 from yeast lysates has no detectable affinity for its syntaxin homologue, and instead binds to assembled SNARE complexes¹⁰. In this case, Sec1 may be acting in conjunction with SNARE complexes to catalyse membrane fusion.

After the target and vesicle membranes have fused, 'cis'-SNARE complexes remain assembled in the fused membrane until they are actively disassembled by a process that requires further proteins and energy supplied by ATP (step 4 in Fig. 1). At this point, the inhibitory domain of syntaxin binds the H3 helix, forming the closed conformation⁵, and so prevents the reassembly of SNARE complexes (step 5). If the interaction between the inhibitory domain and the H3 helix is too weak, other factors may be required at this step to hold syntaxin closed. Misura *et al.* propose that nSec1 is needed here, to bind to the closed conformation of syntaxin and so prevent reassembly of *cis*-SNARE complexes. But, as Misura *et al.* point out, a strictly inhibitory role for Sec1 proteins is inconsistent with genetic data from several organisms indicating that Sec1 function is essential to activate vesicle fusion with membranes¹¹.

Isolation of two classes of Sec1 mutant from *Drosophila*¹² also shows that Sec1 may have a dual function — inhibition and activation — in neurotransmitter release. The inhibitory task of nSec1 in neurons may be to keep syntaxin in a closed conformation, if neuronal syntaxin alone prefers the open form. On binding other factors, nSec1 may

Cell biology

Rafting vesicles

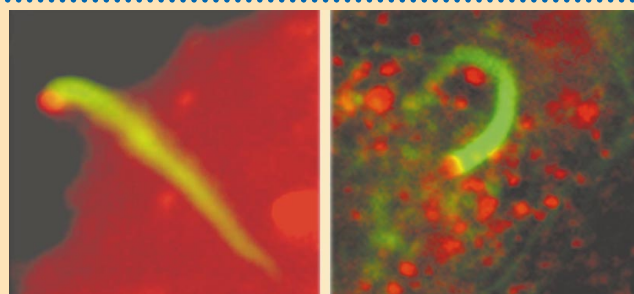
Just as organisms move, so too do their cells and even the organelles within them. For example, vesicles — intracellular transport vehicles — move from organelle to organelle to reach the plasma membrane that encloses a cell, or in the reverse direction.

And just as animals need their skeletons for movement, there's a suspicion that the movement of vesicles relies in some way on the network of proteins that make up a cell's intracellular skeleton — the cytoskeleton. Indeed, some pathogens, such as *Listeria monocytogenes*, take over the cytoskeletal protein actin to propel themselves around the cell. So it seemed plausible that the cell itself makes use of its skeleton for vesicle movement. Now, however,

A. L. Rozelle and colleagues offer evidence for a direct linkage between vesicle transport and actin (*Current Biology* **10**, 311–320; 2000).

Rozelle *et al.* started by looking at a lipid molecule called phosphatidylinositol-4,5-bisphosphate (or PIP₂ for short), which regulates both cytoskeletal and vesicle-trafficking proteins. An enzyme that results in the production of PIP₂ led to the formation of actin-containing 'comet tails' in mouse fibroblast cells. In the pictures reproduced here, green staining identifies these comets, and the red circular shapes at the ends of the comets are vesicles.

The authors also found that comet formation relies on the cell recruiting certain effector proteins to the vesicles found at the head of



comet tails. So, PIP₂ and these proteins may work together in some way to generate comets — probably, Rozelle *et al.* suggest, through two well-known actin-polymerizing proteins.

Rozelle *et al.* next showed that the vesicles at the heads of the comets were derived from an organelle called the Golgi complex or from the plasma membrane. These and other cellular membranes contain microregions, known as 'rafts' because cholesterol and

sphingolipids 'float' in tight-knit groups in these areas. The vesicles at the comet heads mainly budded off from these rafts.

It seems certain that cells do use their cytoskeleton to move vesicles about. But what exactly does actin do? It might help the vesicle-budding process, or it could be a road along which motor proteins carry vesicles. Or, as suggested by Rozelle *et al.*, the comets themselves might push the vesicles along. **Amanda Tromans**

Daedalus

Natural cunning

Nature's composite materials are far superior to our own. Bone (collagen and apatite), wood (cellulose and lignin) and mother of pearl (protein and calcite) all combine strength with enviable toughness. Mineral-filled polymers are far inferior.

The secret, of course, is in the microstructure. The hard particles of a natural composite have a closely defined shape and packing. And the 'glue' holding them together has an ideal tenacity. Chemists are quite good at crystallizing fine particles in desired shapes, such as needles or platelets. But how to pack such particles properly in an adhesive binder?

Daedalus has a novel answer. He wants to pack the particles dry, by vibration and compaction, possibly in an electric field to encourage their orientation, and add the organic component later. It will be a gaseous monomer — ethylene or propylene are the obvious choices for pilot work. Its low viscosity, independent of pressure, will let it percolate freely between the particles even under enormous pressure. Cunningly, Daedalus will then initiate polymerization with X-rays. The gas will hardly absorb them at all. But the mineral particles will block them, liberating photo-electrons copiously on their surfaces. So polymer chains, initiated by and anchored to these surface sites, will start to grow out into the monomer. Many of them will hit another particle, and will terminate on and anchor to it, binding the two particles together by a strong polymer chain and giving ideal adhesion.

But adhesion is only half the problem. A tough composite must be flexible as well. Fortunately, a polymer chain grows in a wandering, self-tangling way. Two particles tied by such a chain could separate a little, untangling the chain slightly. When the load came off, the chain would re-tangle, hauling them together again.

Daedalus's new composites will challenge bone, ivory, jade and mother of pearl on their own ground. Light and tough, they will be moulded like Bakelite, or tooled into shape from billets. They will make wonderfully robust dishes, handles, ornaments — all the small change of material life. And with an elasticity and tenacity so similar to bone and dentine, they will be ideal for false teeth, hip joints, and implants of all kinds. **David Jones**

The Further Inventions of Daedalus (Oxford University Press), 148 past Daedalus columns expanded and illustrated, is now on sale. Special *Nature* offer: m.curtis@nature.com

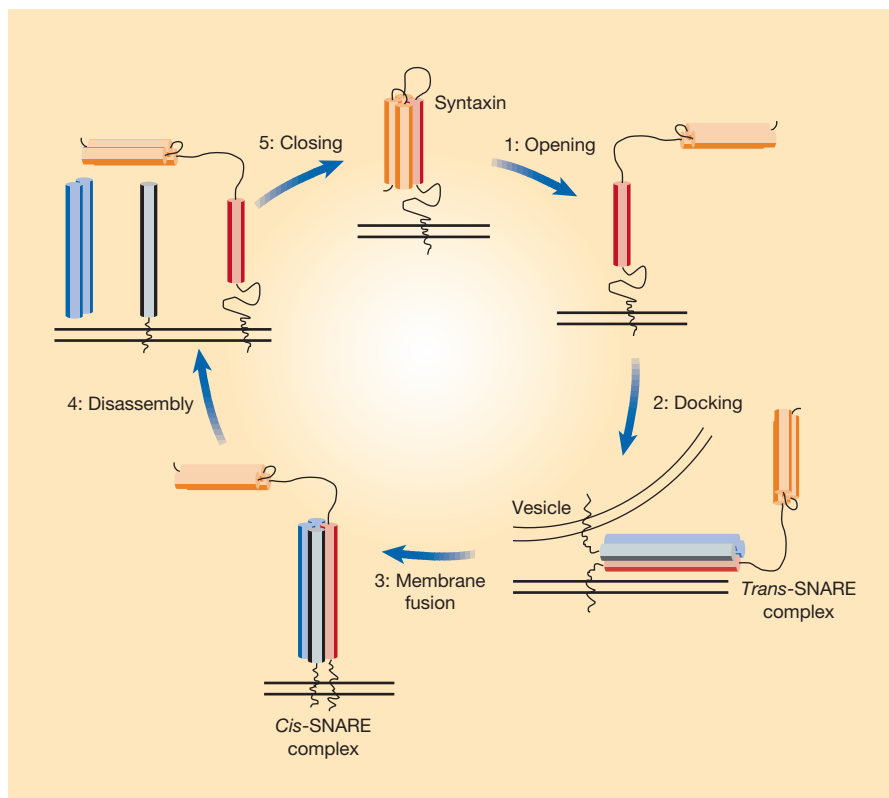


Figure 1 A simplified model for membrane fusion, showing SNARE-complex assembly and conformational changes in syntaxin. Step 1, opening. Syntaxin, a SNARE protein on the target membrane, undergoes a structural change from a 'closed' to an 'open' form. In the open form, the inhibitory domain (orange) is dislodged from the H3 helix (red). Step 2, docking. The vesicle bears a vesicle-SNARE (grey), which assembles with syntaxin's H3 helix and another target-membrane SNARE (blue) to form a 'trans'-SNARE complex, pinning the vesicle to the target membrane. Step 3, membrane fusion. After membrane fusion, the vesicle- and target-membrane SNAREs span the same membrane in a 'cis'-SNARE complex. Step 4, disassembly. SNAREs are separated and recycled for further rounds of fusion. Step 5, closing. The inhibitory domain of syntaxin binds to its H3 helix to form the closed conformation. In this form, syntaxin cannot assemble into a SNARE complex. Sec1 (not shown here) is essential for membrane fusion, but the point at which it acts is unknown. Misura *et al.*¹, from their structure of neuronal Sec1 (nSec1) and syntaxin 1a, propose that nSec1 holds syntaxin 1a in an inactive conformation, and that the binding of other factors to nSec1 may induce it to change conformation itself, so releasing syntaxin 1a in an open form. nSec1 bound to syntaxin may also prevent the reassembly of cis-SNARE complexes after their disassembly.

change its conformation and release open syntaxin 1a, as described above. Yeast syntaxin, however, may not require an extra inhibitor if its inhibitory domain is sufficient to favour the closed conformation. For this reason, yeast Sec1 may not pair with syntaxin as an inhibitor. Consistent with these ideas, SNARE-complex assembly is limited by the rate of syntaxin opening in both yeast and neurons; this rate appears to be at least an order of magnitude slower for the yeast syntaxin homologue⁵ than for its neuronal counterpart⁴ *in vitro*.

Although the details of the interaction between nSec1 and syntaxin 1a are now clear, the purpose of this partnership has not been resolved. Sec1's function in vesicle fusion has yet to be defined biochemically. But for now, ideas about Sec1 should take into consideration the evidence that Sec1 proteins can bind to the four-helix bundle of syntaxin and perhaps also to the four-helix bundle of an

assembled SNARE complex. In the end, syntaxin may not be the only one that is changing dance partners. ■

Chavella M. Carr and Peter J. Novick are at the Yale University School of Medicine, Department of Cell Biology, 333 Cedar Street, PO Box 208002, New Haven, Connecticut 06520-8002, USA. e-mail: peter.novick@yale.edu

- Misura, K. M. S., Scheller, R. H. & Weis, W. I. *Nature* **404**, 355–362 (2000).
- Sutton, R. B., Fasshauer, D., Jahn, R. & Brünger, A. T. *Nature* **395**, 347–353 (1998).
- Fernandez, I. *et al.* *Cell* **94**, 841–849 (1998).
- Hanson, P. I., Otto, H., Barton, N. & Jahn, R. *J. Biol. Chem.* **270**, 16955–16961 (1997).
- Nicholson, K. L. *et al.* *Nature Struct. Biol.* **5**, 793–802 (1998).
- Söllner, T. *et al.* *Nature* **362**, 318–324 (1993).
- Weber, T. *et al.* *Cell* **92**, 759–772 (1998).
- Hata, Y., Slaughter, C. A. & Südhof, T. C. *Nature* **366**, 347–366 (1993).
- Pevsner, J. *et al.* *Neuron* **13**, 353–361 (1994).
- Carr, C. M., Grote, E., Munson, M., Hughson, F. M. & Novick, P. J. *J. Cell. Biol.* **146**, 333–344 (1999).
- Halachmi, N. & Lev, S. *J. Neurochem.* **66**, 889–897 (1996).
- Wu, M. N., Littleton, J. T., Bhat, M. A., Prokop, A. & Bellen, H. J. *EMBO J.* **17**, 127–139 (1998).

Obituary

Robert Rathbun Wilson (1914–2000)

When visiting Fermilab, the site of the massive synchrotron at Batavia in Illinois, one passes through a living memorial to one of the great men of physics — Robert R. Wilson, who died on 16 January at the age of 85. His art, architecture and love of the environment, and his integration of the machines of high-energy physics with nature and with the scientists that use them, are evident everywhere. The Fermilab site represents the ultimate synthesis of his many talents. Wilson's death marks the end of an era when a few giants led us down a path lined with Nobel prizes.

Scientifically, Wilson's greatest contribution was in developing and building accelerators of ever-increasing energy. An accelerator furnishes a 'microscope' for looking inside protons, and the higher the energy of the machine, the finer the detail that can be seen. Wilson's early work as a graduate student at Berkeley with E. O. Lawrence involved cyclotrons working at 20 megaelectron volts (MeV) and his final machine at Fermilab, the Tevatron, is working at 1 TeV (10^{12} eV). The fruits of his drive for higher energy have been the discoveries of the two heaviest quarks, the b and top quarks, at the Tevatron. This facility is used by thousands of physicists and will continue to be the highest energy machine until the Large Hadron Collider comes on line at CERN.

Wilson's love of accelerators began at Cornell University in 1947. Several 300-MeV electron synchrotrons were being built by universities; one was at Cornell. The energy was both fortunate and unfortunate. The threshold for making the newly discovered pi mesons using an electron synchrotron was 150 MeV. These mesons form the glue that holds nuclei together, and there was even an indication that gamma rays could excite them to resonance levels within the proton. The idea that the proton had a resonance was pretty heady stuff — it meant that the proton had structure! The proof required a machine that could go to energies above 300 MeV. At Caltech we had a new machine that could go to 500 MeV, constructed from the quarter-scale model of the Bevatron that had been given to us by E. O. Lawrence. So we were ahead for a while (but not for long!).

In 1952, E. Courant, M. S. Livingston and H. Snyder invented a technique for strong focusing of particle beams. Wilson saw that this discovery promised higher



Artist, architect, humanitarian and physicist

energy and cheaper machines, because the magnets would have a much smaller aperture. He built the first working model of a strong-focusing synchrotron, which verified the theory and showed the promise of this new technology. Wilson continued to push accelerator technology at Cornell, and as a result the Cornell Laboratory of Nuclear Studies continues to be a unique university-centred high-energy physics laboratory.

Wilson's pioneering work at Cornell made him the ideal leader in 1967 for the newly proposed proton synchrotron to be built at Batavia, Illinois. This project was not without controversy. There was a more conventional design for the machine and it was suggested that Wilson had no experience with the more complicated proton accelerators. Nevertheless, he boldly proposed an innovative design. Although the design energy was 300 GeV, I'm sure that it never entered his mind that this would be the final energy. He succeeded in building the original machine and the laboratory to support it, and still had \$10 million left. Moreover, the machine routinely operated at 400 GeV and formed the basis for the present Tevatron.

The money left over, and the use of superconducting technology to build magnets with very high fields, held the promise that a 1-TeV machine could be built in the same tunnel. Whether Wilson

or the government owned the money was contested by some, but not by Bob! He put together a strong team and started to develop suitable magnets. A 10-foot-long magnet was constructed and failed miserably. In typical fashion, he immediately started a model magnet programme, building one-foot-long magnets with the goal of learning how to tame this new technology. This programme allowed us to build model magnets embodying new concepts in just a few weeks. At the same time, a large plant had to be built to supply liquid helium to the six-kilometre circumference of the ring, and industry had to be encouraged to supply tons of superconductor instead of a few pounds. It was Wilson's single-minded pursuit of this dream that finally resulted in the Tevatron and the present pre-eminent position of Fermilab.

Bob Wilson was much more than just an accelerator builder. He designed the whole laboratory, using his skill as an artist and sculptor to carve a unique campus from the flat cornfields of Illinois. The central high-rise building was inspired by the cathedral at Beauvais; the farm houses were saved and used to house scientists; the prairie was restored; trees were planted and sculpture was installed; and innovative architecture was used everywhere. This was to be a laboratory that would attract citizens as well as scientists. Wilson Hall is not only the highest point around, but anyone can go to the fifteenth floor and look out at Illinois as well as receiving an explanation of the science that goes on at Fermilab. Wilson's concern for human rights was manifested in his equal-opportunity stance when construction began.

Wilson made his mistakes, of course. Indeed, the laboratory has many examples of failed experiments. But experiments teach, whether they are a success or not. The model magnet with which he is pictured here failed, but the Tevatron is a tremendous success. Wilson Hall is being repaired because temperature stress has induced fractures in some of the concrete elements. But the early cathedral builders who inspired Wilson also experimented and had failures. The spire of Beauvais cathedral collapsed, suffering a fate far worse than the minor troubles at Fermilab. Jean Mignot, one of the master masons for Milan Cathedral, is quoted as saying "*ars sine scientia nihil est*" ("art without science is nothing"). Bob Wilson certainly believed this — and that science without art is nothing.

Alvin V. Tollestrup

Alvin V. Tollestrup is at Fermilab, PO Box 500, Batavia, Illinois 60510-0500, USA.
e-mail: alvin@fnal.gov

Y-chromosome variation and Irish origins

A pre-neolithic gene gradation starts in the near East and culminates in western Ireland.

Ireland's position on the western edge of Europe suggests that the genetics of its population should have been relatively undisturbed by the demographic movements that have shaped variation on the mainland. We have typed 221 Y chromosomes from Irish males for seven (slowly evolving) biallelic and six (quickly evolving) simple tandem-repeat markers. When these samples are partitioned by surname, we find significant differences in genetic frequency between those of Irish Gaelic and of foreign origin, and also between those of eastern and western Irish origin. Connaught, the westernmost Irish province, lies at the geographical and genetic extreme of a Europe-wide cline.

Surnames have been used in Ireland from about AD 950 as markers of complex local kinship systems. As both surnames and Y chromosomes are paternally inherited, we divided our Irish sample into seven surname cohorts for which ancient geographical information is known, with some error. Four are of prehistoric, Gaelic origin (Ulster, Munster, Leinster and Connaught) and three are diagnostic of historical influx (Scottish, Norman/Norse and English)¹.

The biallelic markers (*SRY-1532*, *M9*, *YAP*, *SRY-2627* (ref. 2); *SRY-8299* (ref. 3); *sY81* (ref. 4); and *92R7* (ref. 5)) define nine haplogroups (clusters of genetic variants) which are highly non-randomly distributed among human populations⁶, including our samples. In particular, haplogroup 1 (hg 1) has a very high frequency in Ireland (78.1% in the island as a whole).

Surname subdivision reveals a cline in Irish samples, with exogenous samples clearly showing lower frequencies (English, 62.5%; Scottish, 52.9%; Norman/Norse, 83.0%) than Gaelic Irish samples (Leinster, 73.3%; Ulster, 81.1%; Munster, 94.6%),

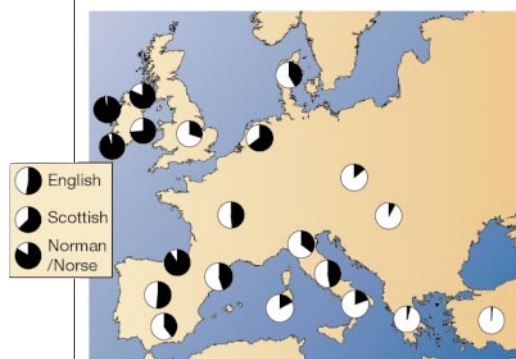


Figure 1 Distribution of observed and estimated haplogroup 1 Y-chromosome haplotypes in Europe. A cline stretches from a frequency of 1.8% in Turkey to peaks in the Basque country (89%) and the west of Ireland (98% in Connaught, the westernmost marker).

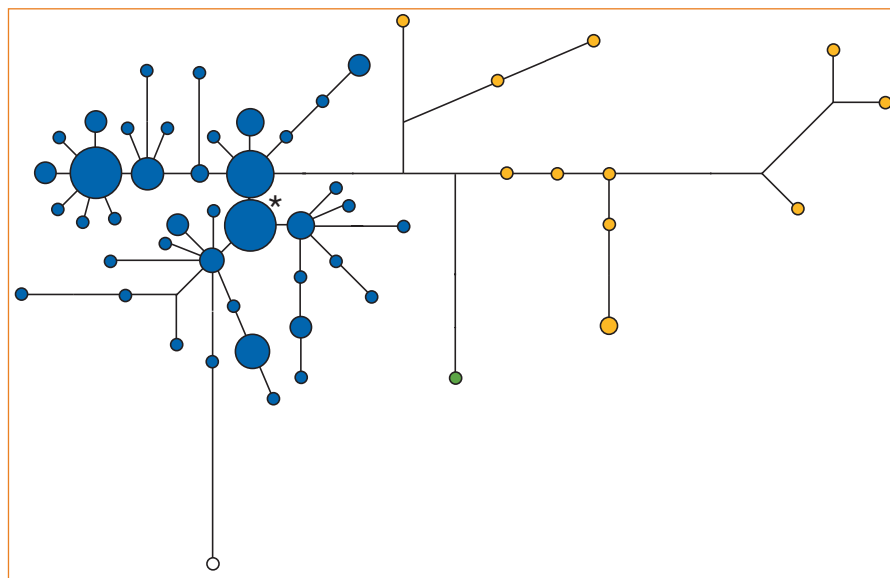


Figure 2 Consensus maximum parsimony networks of Irish Gaelic haplotypes summarizing both single tandem-repeat (*DYS19*, *DYS3891*, *DYS390*, *DYS391*, *DYS392* and *DYS393*; ref. 14) and biallelic variation. Branch lengths are proportional to the number of mutational steps and node areas are proportional to haplotype frequencies. Haplogroup (hg) 1, blue; hg 2, yellow; hg 21, green; hg 26, white. Asterisk, estimated ancestral haplotype. The separate clustering of haplogroups and the tight clustering of the Irish hg 1 haplotypes around a few numerous, central variants are constant through all most-parsimonious trees and were resistant to repetition of the analysis with randomized inputs.

which almost reach fixation in the westernmost province (Connaught, 98.3%). These highly significant differences in the frequency of hg 1 between Irish Gaelic and non-Gaelic Y chromosomes ($P < 0.001$) and between eastern and western Gaelic Y chromosomes ($P < 0.001$) persist when duplicated surnames are removed.

Eighty per cent ($n = 26$; ref. 7) of European hg 1 Y chromosomes belong to 'haplotype 15', defined by using the complex p49f/*TaqI* polymorphic system⁸. Using this relationship, we estimated that hg 1 frequencies follow a cline within Europe⁹, extending from the Near East (1.8% in Turkey) to a peak in the Spanish Basque country (89%; ref. 10) in the west (Fig. 1). This cline mirrors other genetic gradients in Europe and is best explained by the migration of Neolithic farmers from the Near East⁹. When the surname-divided Irish data are appended to this cline, it continues to the western edge of Europe, with hg 1 — the putative pre-Neolithic western European variant — reaching its highest frequency in Connaught (98.3%).

In a maximum-parsimony phylogenetic analysis of both biallelic and simple tandem-repeat (STR) variation between Irish Gaelic haplotypes (Fig. 2), the hg 1 chromosomes cluster together tightly, with the highest-frequency haplotypes occupying central positions, suggesting a coherent common ancestry. The smaller number of

non-hg 1 haplotypes shows no such coherence, consistent with their being immigrants. Their concentration in the eastern Gaelic cohorts may be indicative of a pre-historic influx or of later gene flow across the linguistic barrier from historical migrant groups.

These findings suggest that hg 1 is the earlier, indigenous Irish variant. By taking the ancestral haplotype as that with the most common allele for each STR and calculating the average squared distance¹¹ (assuming a generation time of 27 years and a mutation rate of 0.21%; ref. 12) between it and all variants (Fig. 2), we estimate a date for Irish hg 1 coalescence of 4,200 BP (95% c.i. 1,800–14,800 BP). This relatively recent date (a global estimate of hg 1 coalescence is 30,000 BP; ref. 13) falls well within Ireland's 9,000-year history of human habitation. Although error margins are considerable and include uncertainty related to method, this also provides an upper bound for any agriculturally facilitated population expansion, which, at the fringe of Europe, may have taken place in an insular Mesolithic population of hg 1 genotype.

Emmeline W. Hill*, Mark A. Jobling†, Daniel G. Bradley*

*Department of Genetics, Trinity College, Dublin 2, Ireland

e-mail: dbradley@mail.tcd.ie

†Department of Genetics, University of Leicester, University Road, Leicester LE1 7RH, UK

- MacLysaght, E. *The Surnames of Ireland* (Irish Academic, Dublin, 1997).
- Hurles, M. E. *et al. Am. J. Hum. Genet.* **63**, 1793–1806 (1998).
- Whitfield, L. S. *et al. Nature* **378**, 379–380 (1995).
- Seielstad, M. T. *et al. Hum. Mol. Genet.* **3**, 2159–2161 (1994).
- Hurles, M. E. *et al. Am. J. Hum. Genet.* **65**, 1437–1448 (1999).
- Jobling, M. A. & Tyler-Smith, C. *Trends Genet.* **11**, 449–456 (1995).

- Jobling, M. A. *Hum. Mol. Genet.* **3**, 107–114 (1994).
- Ngo, K. Y. *et al. Am. J. Hum. Genet.* **38**, 407–418 (1986).
- Semino, O. *et al. Am. J. Hum. Genet.* **59**, 964–968 (1996).
- Lucotte, G. & Hazout, S. *J. Mol. Evol.* **42**, 472–475 (1996).
- Goldstein, D. B. *et al. Proc. Natl Acad. Sci. USA* **92**, 6723–6727 (1995).
- Heyer, E. *et al. Hum. Mol. Genet.* **6**, 799–803 (1997).
- Hammer, M. F. *et al. Mol. Biol. Evol.* **15**, 427–441 (1998).
- Kayser, M. *et al. Int. J. Legal Med.* **110**, 125–133 (1997).

Critical phenomena

Fluctuations caught in the act

When a system approaches a critical point, strong fluctuations develop on every scale, from molecules to the entire system. Here we show that critical fluctuations in the domains in a lipid monolayer can be captured and measured by immobilizing it on a solid support and visualizing the pattern using atomic-force microscopy. Such fluctuations in physical properties may reflect the nanometre-scale domain organization in the lipid-bilayer component of biological membranes.

Large numbers of interacting particles show dramatic phase transitions and critical phenomena. At a critical point, molecular-scale interactions build up to macro-

scopic scales, resulting in universal behaviour¹. The discovery of critical opalescence at the critical point of carbon dioxide², caused by density fluctuations scattering light in the fluid, stimulated attempts to glimpse these fluctuations at all scales.

A monolayer of phospholipids at the interface of air and water can be brought close to a critical point by varying the temperature and surface pressure of the monolayer in a Langmuir trough³. We immobilized the monolayer by transferring it, by horizontal dipping, to a solid, hydrophilic substrate of mica using Langmuir–Blodgett techniques⁴. We then imaged critical fluctuations varying in size from nanometres to micrometres using contact-mode atomic-force microscopy.

Figure 1a and b shows two different lipid monolayers — dimyristoyl phosphatidylcholine (DMPC) and dipalmitoyl phosphatidylcholine (DPPC) — close to their critical points. When the monolayer passes through the critical point, the acyl chains of the lipids become disordered and the monolayer becomes thinner. As atomic force microscopy maps the height contours of the monolayer, these images show a monolayer structure with domains of one phase within the other.

The height difference between the two types of domain corresponds to the chain-length difference between an ordered and a disordered lipid acyl chain. Near the critical point, these domains are dynamic and fluctuate strongly when the monolayer is at the air–water interface; we capture them immobilized during the transfer process. The morphology of the domains is very ramified, owing to the vanishing line tension near criticality. Domains of all sizes appear, indicating that there is no characteristic length scale — the hallmark of critical fluctuations.

Figure 1c shows a quantitative analysis of the domain patterns in terms of the structure factor, $S(q)$, where q is the length of a two-dimensional wavevector. The structure-factor data for both DMPC and DPPC scale as a power law, $S(q) \sim q^{-2x}$, with $x \approx 1$. Within the accuracy of the data, this is consistent with the critical-point behaviour associated with the universality class of the two-dimensional Ising model¹.

Lipid domains in monolayers made of phospholipid mixtures can be immobilized

and imaged using similar techniques. Mixtures of phospholipids with acyl chains of different lengths show compositional demixing on the nanometre scale⁵.

Monolayers are simple model systems of the lipid bilayer component of cell membranes⁶. Functions supported by lipid bilayers, such as phospholipase activity⁷ and protein binding⁸, are correlated with lipid-domain formation in the nanometre range that is controlled by the underlying phase transition in the bilayer. Biological membranes contain many different lipid species that are probably organized heterogeneously in the form of domains^{9–11} or ‘rafts’¹² in the plane of the membrane, though the principles underlying this organization remain largely unknown. Our results suggest that fluctuations in lipid bilayer properties, for example, density or composition, may be responsible for lipid-domain formation in biomembranes.

Lars K. Nielsen*, Thomas Bjørnholm†, Ole G. Mouritsen*

*Department of Chemistry, Technical University of Denmark, DK-2800 Lyngby, Denmark

†CISMI, Laboratory for Materials Science, Department of Chemistry, University of Copenhagen, Fruebjergvej 3, DK-2100 Copenhagen, Denmark

e-mail: ogm@kemi.dtu.dk

- Stanley, H. E. *Rev. Mod. Phys.* **71**, S358–S366 (1999).
- Andrews, T. *Phil. Trans. R. Soc. Lond.* **159**, 575–590 (1869).
- Albrecht, O., Gruler, H. & Sackmann, E. *J. Phys. (Paris)* **39**, 301–313 (1978).
- Knobler, C. M. & Schwartz, D. K. *Curr. Opin. Colloid Interface Sci.* **4**, 46–51 (1999).
- Nielsen, L. K., Bjørnholm, T. & Mouritsen, O. G. *J. Phys. Cond. Mat.* **12**, 309–314 (2000).
- Mouritsen, O. G. & Andersen, O. S. (eds) *Biol. Skr. Kgl. Dan. Vid. Selsk.* **49**, 1–214 (1998).
- Hønger, T. *et al. Biochemistry* **35**, 9003–9006 (1996).
- Mustonen, P. *et al. Biochemistry* **26**, 2991–2997 (1987).
- Mouritsen, O. G. & Jørgensen, K. *Curr. Opin. Struct. Biol.* **7**, 518–527 (1997).
- Edidin, M. *Curr. Opin. Struct. Biol.* **7**, 528–532 (1997).
- Korlach, J., Schwillie, P., Webb, W. W. & Feigenson, G. W. *Proc. Natl Acad. Sci. USA* **96**, 8461–8466 (1999).
- Simons, K. & Ikonen, E. *Nature* **387**, 569–572 (1997).

Alzheimer's disease

Apolipoprotein E and cognitive performance

Key proteins implicated in the development of Alzheimer's disease are the β -amyloid precursor protein, which gives rise to the β -amyloid peptides that accumulate in the deteriorating brain^{1,2}, and the different isoforms of apolipoprotein E (apoE). The apoE4 variant increases the risk of developing the disease compared with apoE3 (ref. 3). We have tested the spatial memory of transgenic mice carrying human forms of these proteins and find that it is impaired in mice with apoE4 but not those with apoE3, even though the levels of β -amyloid in their brains are comparable. The fact that apoE3, but not apoE4,

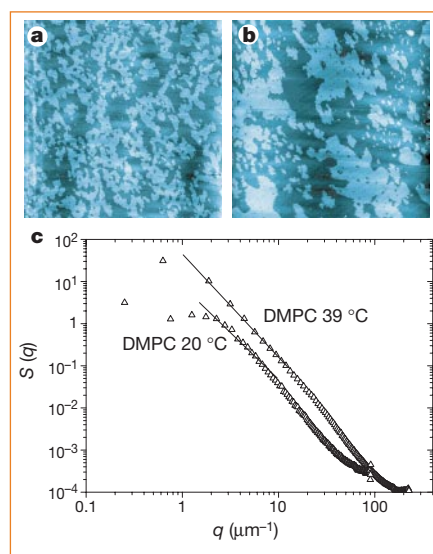


Figure 1 Critical fluctuations in phospholipid monolayers imaged by atomic-force microscopy as a height-difference map. **a, b**, Images of dimyristoyl phosphatidylcholine ($25 \times 25 \mu\text{m}^2$) and dipalmitoyl phosphatidylcholine ($20 \times 20 \mu\text{m}^2$) monolayers at their respective critical points. The monolayers have been transferred from an air–water interface to solid mica supports. The patterns correspond to lipid domains of one phase immersed into the other. The height difference between the light and dark areas is about $5 \text{ \AA}$. **c**, Quantitative analysis of many images like those in **a** and **b** in terms of the structure factor, $S(q)$, in arbitrary units, shown in a double-logarithmic plot as a function of the wavevector, q . The wavevector is related to distance, r , as $q = 2\pi/r$. The data for both phospholipids scale as $S(q) \approx q^{-2x}$, with $x \approx 1$ for $q \leq 20 \mu\text{m}^{-1}$. At low q there is a deviation from this power law because of the finite size of the images analysed.

can protect against cognitive deficits induced by β -amyloid may explain why human apoE4 carriers are at greater risk of developing Alzheimer's than apoE3 carriers.

Transgenic mice that neuronally express human β -amyloid precursor proteins (hAPPs) carrying mutations associated with a familial autosomal dominant form of Alzheimer's develop brain alterations very like those in affected humans^{4–6}. We used mice carrying two such mutations to maximize production of the 42-amino-acid variant of β -amyloid ($A\beta_{1-42}$) that is produced in excess in familial¹ and early sporadic² Alzheimer's. These mice were crossed with mice lacking murine apoE proteins but which express human apoE3 or apoE4 at comparable levels in the brain^{7,8}.

To investigate the effects of hAPP/ $A\beta$ and apoE isoforms on spatial learning and memory, we compared six types of mouse ($n = 14–36$ per genotype) in a water maze test (Fig. 1): bigenic mice (carrying hAPP and apoE3, or hAPP and apoE4), singly transgenic mice (carrying hAPP or apoE3 or apoE4), and controls without human transgenes. At 6 months of age, all groups of male and female mice learned to locate the visible and hidden platform, but there was a clear difference in performance when the platform was removed (probe trial). This probe trial tests whether the mouse spends more time searching in the quadrant that used to contain the platform (target quadrant) than in the other quadrants of the pool, giving a putative measure for retention of spatial memory.

Male hAPP and hAPP/apoE4 mice failed to spend significantly more time in the target quadrant than in the other quadrants, whereas all other male groups showed good memory retention, including hAPP/apoE3 mice (Fig. 1b). Results in 6-month-old female mice were comparable except that singly transgenic apoE4 mice also failed to favour the target quadrant (not shown). The susceptibility of females to apoE4-induced deficits was more striking in aged female mice in which apoE4 alone (without hAPP/ $A\beta$) impaired both spatial memory retention (Fig. 1c) and learning to locate the hidden platform (Fig. 1d). In a more challenging water maze, the learning deficit in apoE4 females was evident at 6 months⁷.

In the absence of hAPP/ $A\beta$ expression, 6- and 18-month-old male apoE4 mice did not differ significantly from age-matched male mice expressing apoE3, mouse apoE, or no apoE at all ($n = 7–27$ mice per group), even in the more difficult water maze test (ref. 7, and data not shown). These sex-dependent differences may relate to the increased susceptibility of women to apoE4-associated cognitive deficits⁷.

Although cerebral amyloid plaques are hallmarks of Alzheimer's disease, they may not cause the associated functional neu-

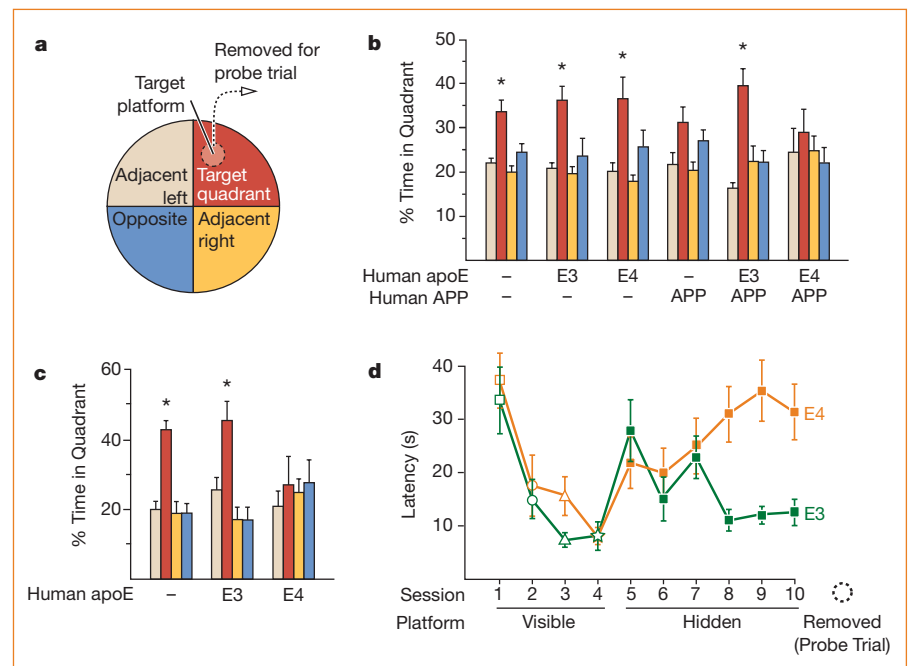


Figure 1 Effects of the expression of human apoE isoforms and hAPP/ $A\beta$ on the spatial memory of mice lacking murine apoE (*ApoE*^{-/-} mice). Endogenous apoE alleles in PDGF-hAPP mice (line J9)⁶ were eliminated in three generations of crosses onto the C57BL/6J-ApoE^{tm1Unc} strain. Heterozygous PDGF-hAPP transgenic *ApoE*^{+/+} mice were then bred with heterozygous NSE-apoE3 or NSE-apoE4 transgenic *ApoE*^{-/-} mice^{7,8}, which had been backcrossed for over ten generations onto a C57BL/6J-ApoE^{tm1Unc} background. Offspring were genotyped^{6,8} and tested at 6 months old. We also assessed 18-month-old NSE-apoE *ApoE*^{-/-} mice^{7,8}. To ensure comparable variability in background genes across groups, we included mice from 13–23 different litters per genotype. Investigators were blinded with respect to genotype. **a**, The water maze (140 cm diameter) was filled with opaque water (24 °C). Mice were trained to locate a visible (days 1–2, location varied) and then a hidden (days 3–5, location constant for each mouse) platform in two daily sessions (3.5 h apart), each consisting of three trials, followed by a probe trial (platform removed), and their performance was monitored by a Noldus Instruments EthoVision video-tracking system. Genotype had no significant effect on fall latency in the rotarod test or on swim speed (not shown), excluding deficits in basic visual and motor skills ($n = 4–30$ mice per genotype, age and gender). **b**, ApoE3, but not apoE4, can prevent hAPP/ $A\beta$ -induced deficits in spatial memory as determined in the probe trial. Results shown are from 6-month-old male mice ($n = 5–27$ per genotype). * $P < 0.05$ (in **b**) or $P < 0.01$ (in **c**) compared with any of the other quadrants by ANOVA and Tukey–Kramer post hoc test. **c**, ApoE4 expression is in itself sufficient to impair retention of spatial memory in female *ApoE*^{-/-} mice ($n = 5–7$ per genotype; age, 18 months). **d**, In aged females, there was a significant difference between the learning curves of apoE3 and apoE4 mice during the hidden platform sessions ($P < 0.05$ for genotype effect and $P < 0.03$ for genotype–learning interaction by multiple-measures ANOVA).

ronal deficit. The intracellular accumulation of β -amyloid or diffusible extracellular forms of this protein may be critical^{1,6,9–13}. Our results indicate that the differential effects of apoE isoforms on hAPP/ $A\beta$ -induced cognitive impairment in 6-month-old hAPP/apoE mice were independent of plaque formation, as no plaques were detectable in hAPP and hAPP/apoE mice at this age (results not shown) by immunostaining⁶. We used an enzyme-linked immunosorbent assay⁴ to compare the levels of diffusible human $A\beta_{1-x}$ (which approximates to the total human $A\beta$) and $A\beta_{1-42}$ in the brains of our transgenic mice, but found no significant difference between hAPP/apoE3 and hAPP/apoE4 mice ($n = 5$ per genotype; means $\pm$ s.d. in nM): hippocampal $A\beta_{1-x}$ was 25.3 ± 7.5 versus 27.8 ± 2.2 , $A\beta_{1-42}$ was 3.4 ± 1.3 versus 4.7 ± 0.8 ; neocortical $A\beta_{1-x}$ was 10.2 ± 3.8 versus 13.5 ± 2.3 , $A\beta_{1-42}$ was 1.6 ± 0.6 versus 2.2 ± 0.4 . We conclude that the difference in the cognitive performance of hAPP/apoE3 and hAPP/apoE4 mice was not due to a variation in the overall amount

of β -amyloid present in their brains.

Compared with apoE3, apoE4 is associated with a greater risk of developing Alzheimer's disease and with a more rapid, age-related cognitive decline even in the absence of frank dementia^{7,8}. Our results indicate that this might be due, at least in part, to the differential capacity of apoE isoforms to inhibit functional neuronal deficits induced by $A\beta$ or other hAPP. It will be important to identify the mechanisms underlying this difference between the two apoE isoforms.

Jacob Raber*†, Derek Wong*, Gui-Qiu Yu*, Manuel Buttini*†, Robert W. Mahley*‡§||, Robert E. Pitas*‡||, Lennart Mucke*†¶

*Gladstone Institute of Neurological Disease, †Cardiovascular Research Institute, Departments of ‡Medicine, †Neurology and ||Pathology, and ¶Neuroscience Program, University of California, San Francisco, California 94141-9100, USA
e-mail: jraber@gladstone.ucsf.edu

- Wilson, C. A., Doms, R. W. & Lee, V. M.-Y. *J. Neuropathol. Exp. Neurol.* **58**, 787–794 (1999).
- Jensen, M. et al. *Ann. Neurol.* **45**, 504–511 (1999).
- Farrer, L. A. et al. *J. Am. Med. Assoc.* **278**, 1349–1356 (1997).

4. Johnson-Wood, K. *et al. Proc. Natl Acad. Sci. USA* **94**, 1550–1555 (1997).
5. Price, D. L. & Sisodia, S. S. *Annu. Rev. Neurosci.* **21**, 479–505 (1998).
6. Hsia, A. *et al. Proc. Natl Acad. Sci. USA* **96**, 3228–3233 (1999).
7. Raber, J. *et al. Proc. Natl Acad. Sci. USA* **95**, 10914–10919 (1998).
8. Buttini, M. *et al. J. Neurosci.* **19**, 4867–4880 (1999).
9. Lue, L.-F. *et al. Am. J. Pathol.* **155**, 853–862 (1999).
10. McLean, C. A. *et al. Ann. Neurol.* **46**, 860–866 (1999).
11. Holcomb, L. A. *et al. Behav. Genet.* **29**, 177–185 (1999).
12. Lambert, M. P. *et al. Proc. Natl Acad. Sci. USA* **95**, 6448–6453 (1998).
13. Hartley, D. M. *et al. J. Neurosci.* **19**, 8876–8884 (1999).

Environmental microbiology

A tetrodotoxin-producing marine pathogen

Disease-related mortalities of sea urchin populations have occurred globally over the past 20 years, although the causative agents have rarely been identified¹. We have discovered a potent new marine pathogen that caused a sudden die-off of the sea urchin *Meoma ventricosa* in Curaçao, Netherlands Antilles, in January of 1997 (ref. 2), and which also has implications for human health. This turns out to be a neurotoxin-producing bacterium that is closely related to the deadly *Pseudoalteromonas haloplanktis tetraodonis*, which is responsible for numerous deaths each year in Japan resulting from the consumption of pufferfish^{3–5}.

The most extensive epizootic reported for sea urchins occurred in the early 1980s, with populations of *Diadema antillarum* being affected throughout the Caribbean sea and western Atlantic: some areas lost more than 97% of mature animals^{6,7}. The ecological impact of the *D. antillarum* die-off on coral reefs was profound, resulting in a dramatic increase of algal cover, which in turn reduced the settlement rates of coral larvae^{7,8}.

M. ventricosa inhabits sandy sediment adjacent to and within sea-grass beds and patch reefs throughout the Florida reef tract and the Caribbean. The die-off in Curaçao extended from the Curaçao harbour to an area 3.5 km downcurrent along the coast². Necropsy of infected animals revealed a progressive loss of spines associated with the disease. Microscopic comparison of the base of the spines and the catch apparatus (where the spine is connected to the skeleton) from affected and healthy animals revealed Gram-negative bacteria in diseased tissue. We isolated these bacteria using a non-selective artificial seawater medium⁹: in all 249 bacterial strains were found in various tissues associated with affected and unaffected urchins.

We exposed these isolates to 95 different carbon sources by using a Biolog microplate assay system^{9,10}. Two metabolic groups (VL-1 and PL-1) were found to occur only in affected tissue (Fig. 1a), so we tested these

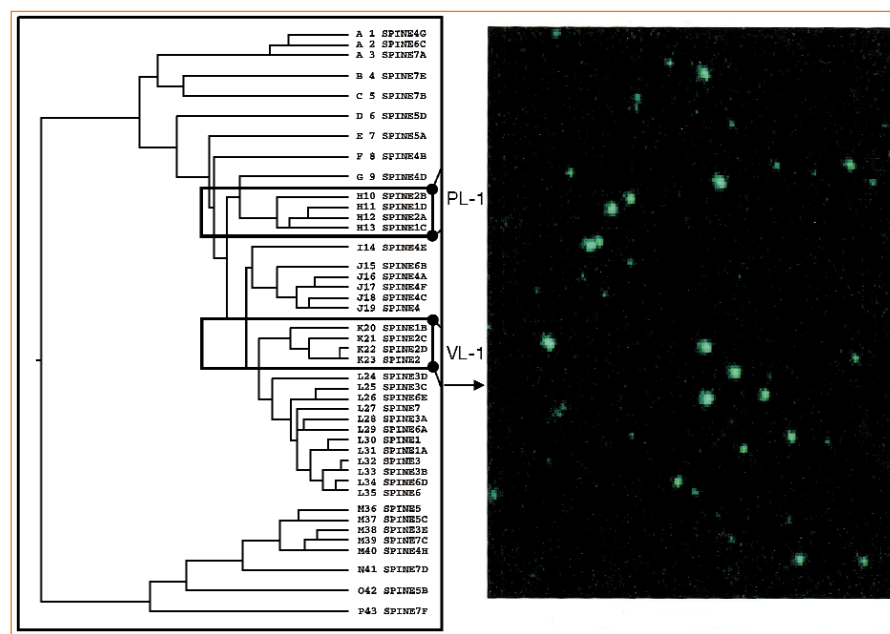


Figure 1 Identification of the new species of tetrodotoxin-producing bacterium. **a**, Dendrogram showing metabolic relationships among bacterial strains isolated from *M. ventricosa* spine catch apparatus. Dendrograms were constructed by clustering hierarchical trees using a nearest-neighbour unweighted pair-group averaging cluster analysis. Isolates PL-1 and VL-1 were indicated as possible pathogens from this grouping. **b**, Immunoassay of the pathogenic isolate VL-1. Fluorescein isothiocyanate-conjugated anti-mouse IgG (green fluorescence) can be seen bound to anti-tetrodotoxin mouse IgG bound to cells producing tetrodotoxin (TTX). Cells were viewed using a confocal laser microscope.

isolates for pathogenicity by infecting the urchin *Lytechinus variegatus*. Aquaria trials with the two isolates showed that urchins inoculated with isolate VL-1 started to shed their spines after two days, whereas control and PL-1-treated urchins remained healthy for over three weeks. We therefore re-isolated VL-1 from the catch apparatus of experimentally infected urchins.

Using DNA-sequence analysis of small ribosomal RNA (16S rDNA), we identified VL-1 as a new bacterium closely related to *Pseudoalteromonas haloplanktis* subsp. *tetraodonis*¹¹, which produces tetrodotoxin² (GenBank accession number, AF154414). Tetrodotoxin has been isolated from various marine organisms, but is best known for its association with pufferfish (order Tetraodontiformes): poisoning by pufferfish (fugu) consumption remains a major health concern, primarily in Japan, where fugu is a delicacy⁴. We tested whether our VL-1 isolate could also produce tetrodotoxin by exposing three-day cultures of VL-1 and PL-1 (as a control) to anti-tetrodotoxin mouse IgG, followed by fluorescein-conjugated anti-mouse IgG¹². We found that a subset of these isolates in the culture could produce tetrodotoxin (Fig. 1b).

To our knowledge, this is the first report of an epizootic affecting *M. ventricosa*. Although the ecological consequences of this urchin die-off are unclear and this epizootic is possibly an isolated incident, it is becoming increasingly important to identify the agents responsible for marine mortality events. In addition, this is the first

time that a tetrodotoxin-producing bacterium has been found to cause disease in a marine animal. And because many urchin species are commercially valuable, our demonstration that this bacterium can cause identical disease symptoms in a different sea urchin species is an important consideration.

Kim B. Ritchie*, Ivan Nagelkerken†, Sara James‡, Garriet W. Smith§

*Department of Biology, University of North Carolina, Chapel Hill, North Carolina 27599-3280, USA

†Carmabi Foundation, PO Box 2090, Piscaderabaai z/n, Curaçao, Netherlands Antilles

‡Department of Microbiology and Molecular Medicine, Clemson University, Clemson, South Carolina 29634-1908, USA

§Biology Department, University of South Carolina, Aiken, South Carolina 29801, USA

e-mail: Smithres@aiken.sc.edu

1. Peters, E. C. in *Pathobiology of Marine and Estuarine Organisms* (eds Couch, J. & Fournie, J.) 393–449 (CRC Press, Boca Raton, 1993).
2. Nagelkerken, I., Smith, G. W., Snelder, E., Karel, M. & James, S. *Dis. Aquat. Org.* **38**, 71–74 (1999).
3. Yotsu, M. *et al. Toxicon* **25**, 225–228 (1987).
4. Noguchi, T. *et al. Mar. Biol.* **94**, 625–630 (1987).
5. Simidu, U., Noguchi, T., Hwang, D. F., Shida, Y. & Hashimoto, K. *Appl. Environ. Microbiol.* **53**, 1714–1715 (1987).
6. Lessios, H. A., Robertson, D. R. & Cubitt, J. D. *Science* **226**, 335–337 (1984).
7. Hughes, T. P. *Science* **265**, 1547–1551 (1994).
8. Lessios, H. A. *Annu. Rev. Ecol. Syst.* **19**, 371–393 (1988).
9. Ritchie, K. B. & Smith, G. W. *Mol. Mar. Biol. Biotechnol.* **4**, 345–352 (1995).
10. Bochner, B. R. *Nature* **339**, 157–158 (1989).
11. Martinez-Murcia, A. J., Benlloch, S. & Collins, M. D. *Int. J. Syst. Bacteriol.* **42**, 412–421 (1992).
12. Brigmon, R. L. & De Ridder, C. *Appl. Environ. Microbiol.* **64**, 3491–3495 (1998).

Three-dimensional structure of the neuronal-Sec1–syntaxin 1a complex

Kira M. S. Misura*, Richard H. Scheller† & William I. Weis*

* Department of Structural Biology, Fairchild Building, Stanford University School of Medicine, 299 Campus Drive West, Stanford, California 94305, USA

† Molecular and Cellular Physiology and the Howard Hughes Medical Institute, Stanford University School of Medicine, Stanford, California 94305, USA

Syntaxin 1a and neuronal Sec1 (nSec1) form an evolutionarily conserved heterodimer that is essential for vesicle trafficking and membrane fusion. The crystal structure of the nSec1–syntaxin 1a complex, determined at 2.6 Å resolution, reveals that major conformational rearrangements occur in syntaxin relative to both the core SNARE complex and isolated syntaxin. We identify regions of the two proteins that seem to determine the binding specificity of particular Sec1 proteins for syntaxin isoforms, which is likely to be important for the fidelity of membrane trafficking. The structure also indicates mechanisms that might couple the action of upstream effector proteins to conformational changes in syntaxin 1a and nSec1 that lead to core complex formation and membrane fusion.

The organization of the membrane compartments of eukaryotic cells and intercellular communication depend on the directed targeting and fusion of transport vesicles with acceptor membranes¹. Docking and fusion of neurotransmitter-containing vesicles with the presynaptic membrane is a well studied example of this process, and the proteins involved are representative of families whose members have equivalent roles in other inter-compartmental transport processes². The formation of a parallel four-helix bundle^{3,4} by proteins present on vesicle and target membranes, designated SNAREs, is thought to contribute to membrane fusion, perhaps by overcoming the energetic barrier to bilayer fusion^{5–7}. In the neuron, the vesicle membrane SNARE VAMP (also known as synaptobrevin) contributes one helix, and the presynaptic membrane SNAREs syntaxin and SNAP-25 contribute one and two helices, respectively, to the complex. After fusion, the complex is dissociated by the action of the ATPase NSF, allowing the SNAREs to be recycled for another round of fusion^{8,9}.

Cytosolic proteins homologous to the yeast SEC1 gene product are also required for vesicle trafficking and membrane fusion. Neuronal Sec1 (nSec1, also known as rbSec1 and munc18-a^{10–12}) binds with nanomolar affinity to syntaxin 1a¹⁰ and forms a complex that is mutually exclusive of the syntaxin 1a–VAMP–SNAP-25 core complex *in vitro*^{13,14}. Overexpression of Rop, the nSec1 orthologue in *Drosophila*, leads to a decrease in neurotransmitter release¹⁵. However, null alleles and certain point mutations of members of the Sec1 family lead to a block in vesicle fusion^{16,17}, which indicates that these proteins not only sequester syntaxins from other proteins but also assist the syntaxins in adopting a functional conformation or facilitate interactions between syntaxins and other proteins. Such chaperone-like activities of Sec1 proteins might protect regions of unstructured polypeptide present in isolated syntaxin; in PC12 cells, syntaxin is unable to exit from the Golgi in the absence of nSec1 (ref. 18). Sec1 family members also interact genetically with Rabs, which are small GTPases that are also essential for vesicle trafficking and membrane fusion¹⁹.

Unlike other syntaxin-binding proteins, nSec1 requires most of the cytoplasmic portion of syntaxin 1a for high-affinity binding²⁰. Syntaxins contain a moderately well conserved amino-terminal domain, known as Habc, whose structure is an antiparallel three-helix bundle²¹. A linker of about 30 amino acids connects Habc to the carboxy-terminal region, designated H3, of the syntaxin cytoplasmic domain. The highly conserved H3 region forms a single, long α -helix when it is part of the core SNARE complex³, but is partly unstructured in isolation^{22,23}. Binding to nSec1 requires both the Habc and H3 regions of syntaxin; conversely, the entire nSec1

protein is required for binding to syntaxin. As only specific combinations of syntaxins and Sec1 family members can interact^{10,24}, Sec1 proteins might also contribute to the specificity of membrane trafficking.

To improve our understanding of the roles of nSec1 and syntaxin 1a in membrane trafficking, we have determined the three-dimensional structure of the complex between these two proteins. The structure reveals major conformational rearrangements of syntaxin 1a relative to the conformations observed in the core SNARE complex and in isolation, and indicates regions of the nSec1 and syntaxin 1a proteins that are likely to determine the specificity between particular isoforms of the two proteins. The structure also indicates mechanisms that might couple the action of Rabs and Rab effector proteins to the release of syntaxin and subsequent core complex formation.

Structure determination

The nSec1–syntaxin complex was produced by mixing lysates from *Escherichia coli* strains expressing either full-length nSec1 (residues 1–594) or the full cytoplasmic domain of syntaxin 1a (residues 1–267), and purifying the complex. All diffraction data were measured from crystals adapted to pH 7.5 in a low-salt buffer to facilitate heavy-atom binding. The structure was solved with one mercury derivative and multiwavelength anomalous dispersion (MAD) phasing by using a crystal containing selenomethionine-substituted nSec1 and native syntaxin 1a protein. The final model, refined against data to 2.6 Å Bragg spacings, consists of residues 4–316, 324–505 and 532–592 from nSec1, residues 27–248 from syntaxin 1a, and 35 water molecules. There is no electron density for the remaining 83 amino acids, which are presumably disordered.

Structure of nSec1

Neuronal Sec1 is an arch-shaped molecule with a central cavity ~ 15 Å wide. The nSec1 polypeptide chain can be divided into three domains (Fig. 1a, b). Domain 1 is formed by residues 4–134 and consists of a five-stranded parallel β -sheet flanked by five α -helices. The order of the strands is 2-1-3-4-5 (Fig. 1a), and is similar in topology to many proteins such as Kex1 and uracil phosphoribosyltransferase. Residues 135–245 and 480–592 comprise the second domain, which like the first domain is an α - β - α fold. The β -sheet of domain 2 features five parallel strands with an additional antiparallel strand on one edge (Fig. 1a). The order of the parallel strands is 1-5-4-3-2 (Fig. 1a), which is different from that in domain 1 and is unique to this protein, indicating that these two domains did not arise from a gene duplication event. In addition, the

connection between strands 2 and 3 ($\beta 8$ and $\beta 9$) is left-handed (Fig. 1a), an unusual feature that might be the result of constraints imposed during folding that orientate the three domains correctly with respect to one other. No electron density is seen for residues 506–531 of nSec1 domain 2, but residues 532–592 form the central two strands of the domain 2 β -sheet and one of the flanking helices. The region between residues 506 and 532 is a site of proteolytic sensitivity that is partly cleaved during protein expression and purification procedures (see Methods).

The third domain, residues 246–479, is a large insertion between the third and fourth parallel strands ($\beta 9$ and $\beta 12$) of domain 2. This domain can be subdivided into a lower half and an upper half (Fig. 1a, b), which we shall refer to here as domains 3a and 3b, respectively. Domain 3b is a cluster of helices that exhibits structural similarity to a variety of helical repeat proteins such as α -karyopherin and Sec17. Domain 3a consists of two long antiparallel helices, a long, twisted β -hairpin that is perched on these helices, and a shorter helix that is orientated perpendicularly to the hairpin and which faces the interior of the protein. Helices $\alpha 12$ and $\alpha 13$ serve as a backbone for domain 3, with $\alpha 12$ continuing into $\alpha 13$, which is the first helix of domain 3b. A single non-helical glycine introduces a 60° bend between these two helices, and might be a hinge between domains 3a and 3b. Domain 3a makes relatively few contacts with any other part of nSec1, suggesting that this collection of secondary structure elements might be able to move with respect to the rest of the protein.

Structure of syntaxin 1a

Residues 27–248 of syntaxin 1a are ordered in this structure, and the

protein is largely helical (Figs 1c, d and 2). The disorder of residues 1–26 and 249–267 is consistent with the results of other structural studies of this protein^{21,23}. The Habc domain forms a three-helix bundle whose structure is essentially the same as this region in isolation; the $C\alpha$ coordinates of the Habc helices superimpose with those of NMR²¹ and crystal (F. M. Hughson, personal communication) structures of this domain with r.m.s. deviations of 1.7 and 0.9 Å, respectively. The Hc helix ends at residue 156 and is nine residues longer in the nSec1–syntaxin 1a complex than in the X-ray and NMR structures of the isolated syntaxin 1a Habc domain. The linker region between Habc and H3, residues 156–187, is ordered as in isolated syntaxin^{22,23}, and residues 162–170 adopt an α -helical conformation. This short helix is on the exterior of the nSec1–syntaxin 1a complex, and is orientated at an angle of $\sim 90^\circ$ with respect to the long syntaxin helices (Figs 1d, 2b). Most of the linker helix side chains are polar, except Leu 165 and Leu 169, which face away from solvent and pack against residues located on the Hc and H3 helices. Although it was possible to trace the main chain path of the entire linker, the assignment of side chains for residues 173–187 was difficult owing to poor electron density; these residues have been modelled as alanine. This feature indicates that there might be multiple conformations in this region of syntaxin 1a.

The C-terminal H3 region consists of residues 188–248 and adopts a structure distinct from that observed in the core complex with VAMP and SNAP-25 (ref. 3). Contacts between H3 and Habc, as well as between H3 and nSec1 (Figs 2, 3b), stabilize H3 in a conformation containing helical regions interrupted with bends and other irregularities. The association of Habc and the N-terminal portion of H3 is consistent with biochemical^{25,26} and NMR^{22,23} data that indicate interactions between these regions in isolated syntaxin. The first 15 residues of H3 form an α -helix (H3a) that interacts with a groove formed by the Hb and Hc helices to create an antiparallel four-helix bundle (Fig. 1c). The hydrophobic H3a residues that face into the centre of the four-helix bundle in the SNARE core complex face Hc and the linker region (Fig. 3b). H3 then bends away from the Habc helices and towards domain 3a of nSec1 (H3b; Figs 1c, d and 2). Although H3b maintains contacts with Hc, the association of H3b with the Habc region does not resemble a canonical four-helix bundle. The initial bending of H3b away from Habc can be attributed to polar residues in the Habc domain of syntaxin 1a that are incompatible with the hydrophobic surface of the H3 helix. Beyond residue 226, which is a conserved glutamine that forms the central layer of the four-helix bundle in the SNARE complex³, the regular secondary structure is interrupted several times. This mixture of turns, a short helix (H3c) and an extended coil fills the central cavity of nSec1 and forms intimate contacts with domains 1

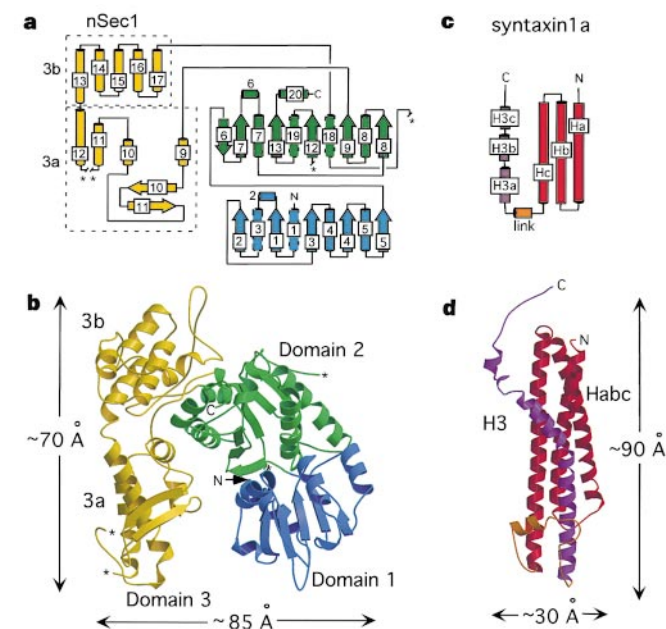


Figure 1 Structure of nSec1 and syntaxin 1a. **a**, Topology diagram of nSec1. Domain 1 is shown in blue, domain 2 in green and domain 3 in yellow. Helices are denoted by cylinders and strands by arrows. Breaks in the structure are indicated with asterisks. In domains 1 and 2, helices with dashed and solid lines lie on opposite sides of the parallel β -sheet. Note the left-handed crossover between strands $\beta 8$ and $\beta 9$. Dashed outlines denote domains 3a and 3b. **b**, Ribbon representation of nSec1, coloured as in **a**. **c**, **d**, Topology (**c**) and ribbon diagrams (**d**) of syntaxin 1a. The Habc domain is shown in red, the Habc/H3 linker in orange and the H3 region in purple. The conformations of nSec1 and syntaxin 1a are as they appear in the protein complex, but have been separated and reorientated here for clarity. In **b** and **d** the approximate dimensions are shown; the dimensions perpendicular to the plane of the page are ~ 45 Å for nSec1 and ~ 35 Å for syntaxin 1a. Panels **b** and **d** were prepared with the program MOLSCRIPT⁴⁷, as were Figs 2 and 5. Secondary structure was assigned by PROCHECK⁴⁸.

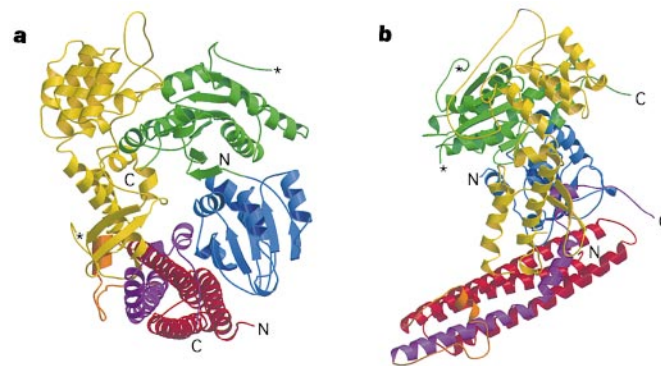


Figure 2 Ribbon representation of the nSec1–syntaxin 1a complex. **a**, View looking down the long syntaxin 1a helices; nSec1 is shown in the same orientation as in Fig. 1b. **b**, As in **a**, but rotated about the vertical axis by 90° . N- and C-termini are indicated, and the colour coding is as in Fig. 1.

and 3a of nSec1 (Fig. 2). NMR and circular dichroism data indicate that the C-terminal half of the H3 region (residues 221–267 in syntaxin 1a and equivalent residues in SSO1) is unstructured in the isolated cytoplasmic portion of syntaxin 1a^{22,23}. The structure of the nSec1–syntaxin 1a complex therefore demonstrates that the H3 region can adopt at least three different conformations: bound to nSec1, bound to VAMP and SNAP-25, and in isolation, where it might exist as a potentially large ensemble of structures.

Interactions between nSec1 and syntaxin 1a

Both the Habc and H3 regions of syntaxin 1a contact nSec1 (Figs 2, 3), which is consistent with earlier deletion mutagenesis studies²⁰. A large (3,932 Å²) amount of solvent-accessible surface area is buried

between the two proteins, although this represents only 10% of the total surface area of the complex. In addition to forming a scaffold that binds the H3 region, the Habc helices contribute approximately half of the residues that contact nSec1. However, the Habc surface area contacting nSec1 is smaller than that of the H3 region (677 Å², or 9% of the total Habc surface area, compared with 1,237 Å² or 27% of the total H3 surface area). The central cavity formed by domains 1 and 3a of nSec1 provides the binding surfaces for syntaxin 1a (Fig. 2). The domain 1 surface (largely formed by residues 38–71; Fig. 3a) contributes more contacts and a greater surface area (1,225 Å²) than domain 3a (725 Å²) to the syntaxin-1a-binding site. Domain 1 contacts syntaxin 1a Habc, H3c and the following C-terminal extended region of H3 (Figs 2, 3). In domain

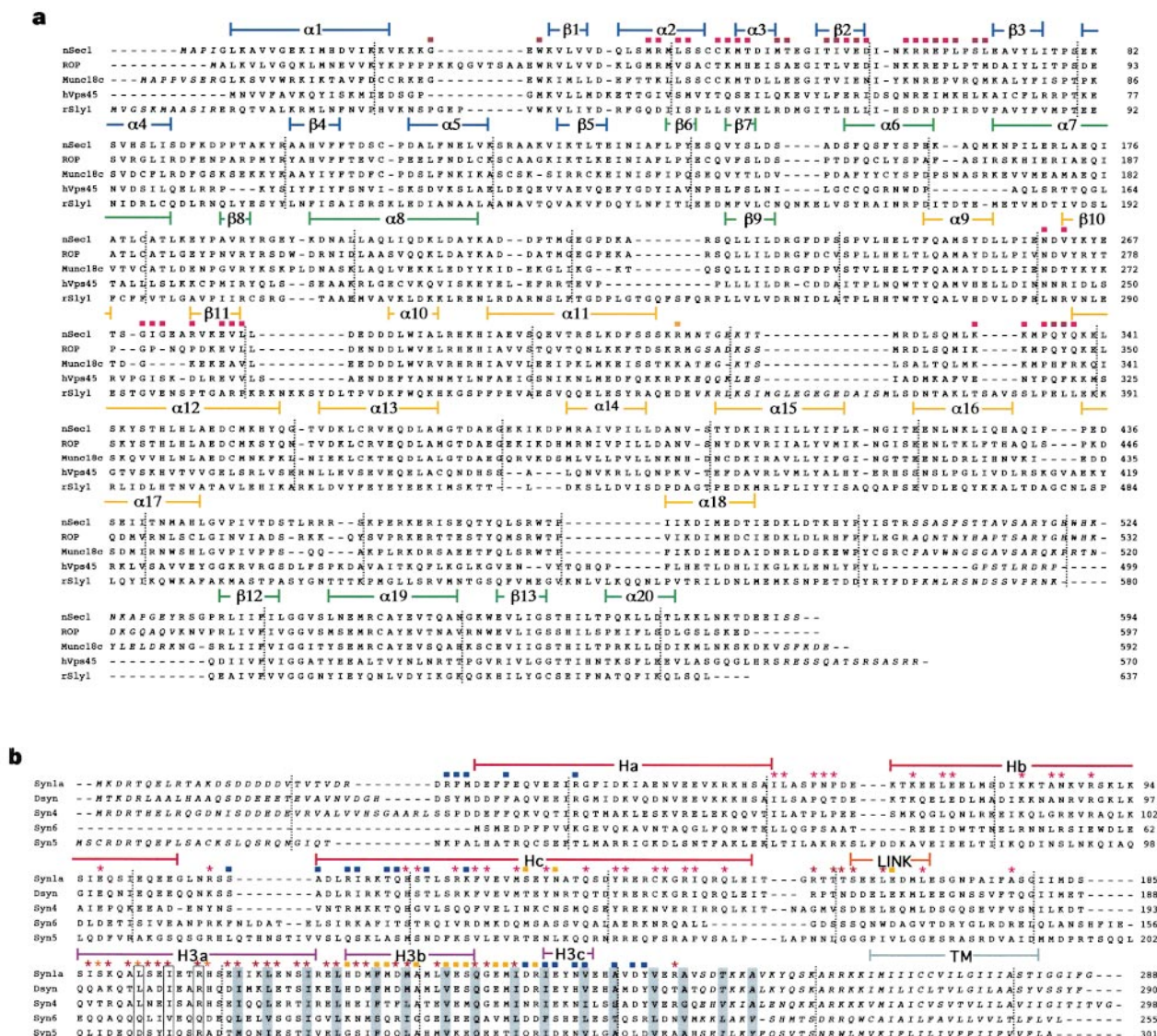


Figure 3 Structure-based sequence alignments. Alignments of four Sec1 family proteins are shown with nSec1 (a) and four syntaxin family proteins with syntaxin 1a (b). nSec1 binds plasma membrane-localized syntaxins 1a, 1b, 2 and 3; Rop binds *Drosophila* syntaxin 1; Munc18c binds the plasma-membrane-localized syntaxin 4; hVps45 binds syntaxin 6 localized in the *trans*-Golgi network and endosomes (ref. 48); and rSly1 binds syntaxin 5 localized in the intermediate compartment and the *cis*-Golgi network. Secondary structure elements are indicated and correspond to the topology diagrams in Fig. 1. The portion of the syntaxin 1a sequence marked LINK corresponds to the short α -helix of the Habc/H3 linker. Vertical dashed lines appear every 20 residues for nSec1 and syntaxin 1a. Italicized residues are not seen in the nSec1–syntaxin 1a complex structure. In a, boxes above certain residues denote an interaction with syntaxin 1a (red with Habc,

orange with the linker and purple with H3). In b, boxes indicate syntaxin 1a residues that interact with nSec1 (blue with domain 1, yellow with domain 3), and asterisks indicate residues that interact with other syntaxin 1a subdomains, coloured as in Fig. 1. H3 residues highlighted in grey are those that form the hydrophobic core of the four-helix bundle in the VAMP–SNAP-25–syntaxin 1a (H3) core complex structure; those outlined in a black box pack against residues of SNAP-25 in the core complex. Certain residues form multiple contacts that could not be represented clearly on this figure, and in these cases only one interaction has been marked. In a, nSec1 residues Met 47, Thr 48 and Val 58 also contact syntaxin Habc; Glu 66 and Gln 336 also contact syntaxin H3. In b, Arg 114 and Asn 135 also contact H3; Arg 125 and Ile 230 also contact nSec1 domain 1; Glu 234 also contacts domain 3a of nSec1.

3a, the inner strand (β 11) of the extended β -hairpin (residues 271–280) contacts the H3b helix as it begins to bend away from the syntaxin Habc domain. Residues 331–338 of domain 3a are located beneath the β -hairpin of nSec1, and also contact the H3b helix (Figs 2, 3).

The nSec1 electrostatic surface potential shows many regions of positive and negative charge (Fig. 4a), whereas syntaxin 1a carries a predominantly negative charge (Fig. 4b). The interacting surfaces of nSec1 and syntaxin 1a are largely polar and complementary in charge. For example, a highly charged cavity in domain 1 (labelled Habc N terminus in Fig. 4) seems to anchor the most N-terminal residues of syntaxin 1a (residues 27 and 28) seen in this structure. A negatively charged cavity with approximate dimensions $20 \text{ \AA} \times 15 \text{ \AA} \times 15 \text{ \AA}$ is found between domains 1 and 2 of nSec1 (Fig. 4a), and might represent a ligand-binding site distinct from that of syntaxin 1a. One face of helix α 2 forms the base of the cavity and the opposite face interacts with syntaxin 1a, so a ligand bound in the cavity might affect the interaction of nSec1 with syntaxin 1a.

Protein kinase C (PKC) modulates exocytosis in several different cell types, possibly by regulation of the nSec1–syntaxin interaction. PKC phosphorylates both Ser 306 and Ser 313 of nSec1, which inhibits formation of the nSec1–syntaxin 1a complex. Once the complex is formed, however, these sites can no longer be phosphorylated²⁷. Ser 313 is located in domain 3a, near the Habc/H3 linker, whereas Ser 306 is farther away from syntaxin 1a. The Habc/H3 linker helix carries a large negative charge (Fig. 4b), and is provided with a favourable environment formed by the nearby positively charged nSec1 domain 3a surface (Fig. 4a). Although a phosphoserine could be accommodated at position 313 without sterically disrupting the syntaxin-binding surface, the introduction of a negative charge would disrupt the electrostatically positive character of domain 3a and the charge complementarity with the syntaxin 1a linker region. Thus the phosphorylation of Ser 313 might promote exocytosis by preventing formation of a high-affinity nSec1–syntaxin 1a complex. The effects, if any, of phosphorylation at Ser 306 are not clear. Cyclin-dependent kinase 5 (CDK5) has also been shown to phosphorylate nSec1, and in combination with accessory proteins might mediate disassembly of the nSec1–syntaxin 1a complex²⁸. The site of CDK5-mediated phosphorylation in nSec1, Thr 574, is located between domains 2 and 3. In the structure observed here, this site would be sterically inaccessible to a protein, indicating that CDK5 interacts with a different conformation of nSec1.

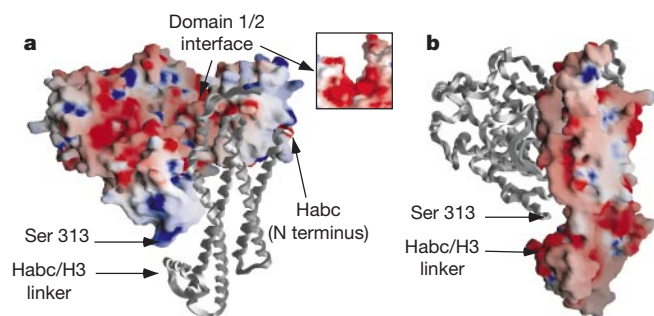


Figure 4 Electrostatic surface potentials of nSec1 and syntaxin 1a. **a**, Electrostatic surface representation of nSec1, with the syntaxin 1a backbone shown as a grey ribbon. The outlined box is an enlarged view of the interface between domains 1 and 2, shown in an orientation rotated 180° about the vertical axis from **a**. **b**, Electrostatic surface map of syntaxin 1a, with nSec1 shown as a grey ribbon. Red patches represent negative potential, blue represents positive potential, and grey or white is neutral. The map is contoured at the $10kT/e$ level. Ser 313 is one of the PKC phosphorylation sites that seems to disrupt nSec1–syntaxin 1a complex formation. The position of the syntaxin 1a Habc/H3 linker is also indicated. **a** and **b** are shown in approximately the same orientation. This figure and Fig. 6 were prepared with GRASP⁴⁹.

Specificity

Most interactions between syntaxin 1a and both surfaces of nSec1 involve main chain–main chain or main chain–side chain contacts that appear to be largely conserved in the Sec1- and syntaxin-like proteins (Fig. 3). A smaller number of interactions are not conserved among Sec1 family members, and might help to determine binding specificity between a particular Sec1-like protein and its partner syntaxin (Fig. 3). One of the most obvious differences among the Sec1-like proteins is the high sequence and size variability in the domain 3a β -hairpin (β 10 and β 11) that interacts with syntaxin. Thus, Munc-18c and Rop have deletions relative to nSec1 that probably result in shorter hairpins, whereas rSly1 contains a nine-residue insertion. In contrast, hVps45 contains the same number of residues and is largely conserved with respect to nSec1 in this region.

The overall conformation of particular syntaxin isoforms, determined in part by residues that do not contact nSec1, is likely to be an important determinant of the specificity between Sec1 and syntaxin family members. For example, syntaxin 4 shares 80% sequence similarity with syntaxin 1a, and forms a stable complex with munc-18c but not with nSec1 (ref. 24). Many of the residues that form contacts between nSec1 and syntaxin 1a are conserved in Munc-18c and syntaxin 4, and therefore are not likely to be determinants of binding specificity. However, certain Habc side chains that contact H3 differ between syntaxins 1a and 4. Similarly, the hydrophobic interaction between the short helix of the Habc–H3 linker structure (Fig. 5c) seems to be conserved in many different syntaxins, but the

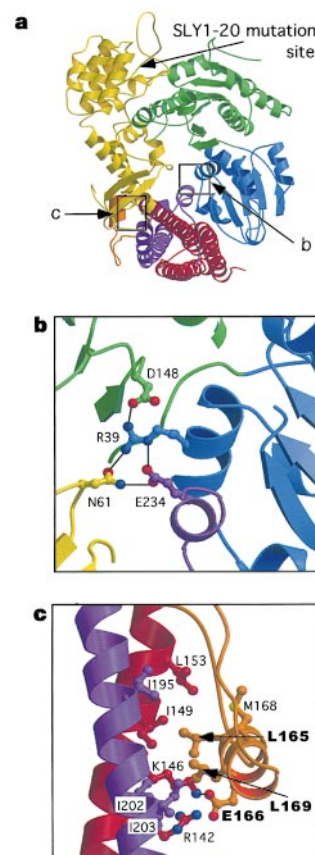


Figure 5 Sites of Sec1 and syntaxin mutations. **a**, Location of the yeast SLY1–20 mutant mapped onto the nSec1–syntaxin 1a complex structure. The orientation is the same as in Fig. 2a. Sites of selected Rop (**b**) and syntaxin 1a (**c**) mutations are outlined in boxes. **b**, View of the Rop Arg 50 → Cys mutation (position 39 in nSec1). **c**, View of the syntaxin 1a Glu 165 → Ala/Leu 166 → Ala mutations in the Habc/H3 linker region. Hydrophobic residues from H3 and Hc which face the linker are also shown in **c**. Residue numbers in all panels correspond to the rat nSec1 and syntaxin 1a sequences.

remainder of the linker region shows considerable variation. For example, Lys 146 in syntaxin 1a Hc forms a salt bridge with Glu 166 of the linker helix, but in syntaxin 4 a valine replaces the lysine, which would probably produce local changes in the packing of these structural elements and their interactions with H3. Differences in contacts between Habc and H3 and in the linker might result in different syntaxins presenting a different H3 conformation to the variable hairpin region in domain 3a of the Sec1-like proteins.

Mutations in Sec1 and syntaxin proteins

The high sequence identity (~65%) between nSec1 and its *Drosophila* orthologue Rop, and between syntaxin 1a and *Drosophila* syntaxin, provides an opportunity to examine the structural effects of Rop mutations. Here we discuss Rop mutants that either decrease (His 302→Tyr, Asp 45→Asn) or increase (Arg 50→Cys, Pro 254→Ser) neurotransmitter release^{29,30}. It is possible that the decrease in transmitter release is due to misfolding of the protein, but it seems that the mutant Rop proteins are expressed in neurons. His 302 of Rop is equivalent to His 293 of nSec1, which is located between domains 2, 3a and 3b. The His→Tyr mutation would alter the packing in this domain interface region. Asp 45 (Asp 34 in nSec1) is positioned between domains 1 and 2, and both carboxylate oxygens form interactions with main-chain amide groups from domain 1. As the amide nitrogen of the asparagine side chain would not be able to form these interactions, the Asp 45→Asn change would alter the structure in this region, which forms part of the syntaxin-binding site. Both of these mutations would cause perturbations in the architecture of the Sec1 protein rather than changing residues that interact directly with syntaxin. In contrast, Arg 50 of Rop (Arg 39 of nSec1) directly contacts Glu 234 (Glu 237 of *Drosophila* syntaxin) in the centre of the protein complex, and also interacts with domain 2 of nSec1 (Fig. 5b). A

cysteine at this position would not be able to perform any of these roles. Pro 254 of Rop corresponds to Pro 242 of nSec1, which is tightly packed between domains 2 and 3b, and a change to serine would alter this packing arrangement. The opposing phenotypes of the two pairs of Rop mutants can be rationalized in terms of a model presented below.

The SLY1–20 mutation, which uncouples Rab function from vesicle transport, is a single amino-acid substitution in the Sec1 family member involved in transport from endoplasmic reticulum to Golgi in yeast³¹. Sequence alignment between the SLY1 protein and nSec1 is difficult, but the position of the SLY1–20 mutation can be assigned to the loop connecting domains 2 and 3b (residues 460–470 in nSec1; Figs 1a, b, 3, 5). This exterior loop is well separated from the syntaxin-binding site, so a Rab or Rab effector protein docking at this site is not likely to interact with syntaxin in the conformation observed here. Instead, the loop might serve as part of a hinge mechanism for releasing syntaxin when acted on by a Rab or a Rab effector. The mutation could allow effector binding in the absence of a Rab protein, restoring wild-type transport activity, or it might permit the release of syntaxin in the absence of an appropriate signal. Interestingly, the loop containing the SLY1–20 mutation lies near the site of the Rop Pro 254→Ser mutation discussed above.

A Leu 165→Ala, Glu 166→Ala double mutation of syntaxin 1a²² abolishes complex formation with nSec1. These residues are in the short linker helix (Fig. 5c), which is separated from nSec1 by ~4.5 Å, so this mutation does not obviously influence direct contacts between syntaxin 1a and nSec1. However, disruption of the leucine hydrophobic contacts and the hydrophilic character of the glutamic acid might destabilize the base of the Habc–H3a four-helix bundle and prevent syntaxin from adopting a conformation capable of binding to nSec1. An Ile 236→Ala mutation of *Drosophila* syntaxin³², and the equivalent change in syntaxin 1a,

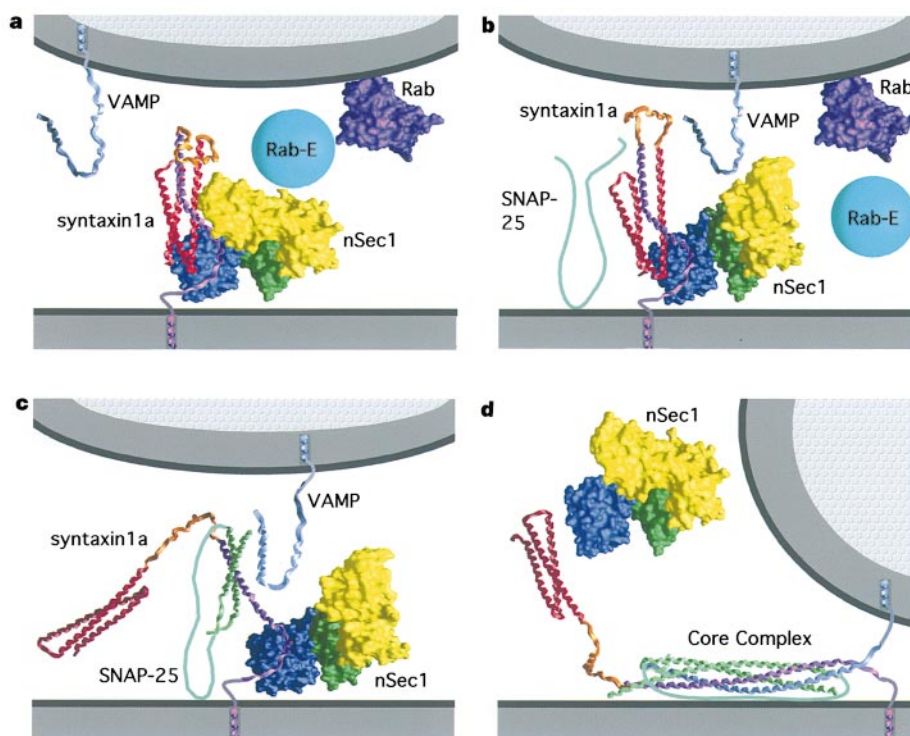


Figure 6 Model for the role of nSec1 in membrane fusion. **a**, nSec1 bound to the conformation of syntaxin 1a observed in the crystal structure. This complex is a recognition site for an arriving vesicle, possibly through a Rab or Rab effector protein. **b**, The action of Rab or Rab effectors or other proteins induces a conformational change in nSec1 and causes the loss of contacts between domain 3a and syntaxin 1a, which destabilizes this region of syntaxin 1a. Exposed residues of H3a interact with SNAP-25,

initiating the formation of the core complex. **c**, The syntaxin Habc domain moves away from the H3 region. The SNAP-25 helices propagate towards the C-terminus, and VAMP begins to bind to the helical H3a region. **d**, The nucleated VAMP–SNAP-25–syntaxin 1a helices assemble rapidly and complete the long straight helical bundle structure (core complex). Formation of the core complex might then promote membrane fusion.

Ile 233→Ala (ref. 20), reduces binding to nSec1. Ile 233 of syntaxin 1a is on the H3c helix, which occupies the central cavity of nSec1 and is close to the position equivalent to the Rop Arg 50→Cys mutation. The Ile side chain participates in extensive hydrophobic packing interactions with domain 1 of nSec1, and a smaller side chain would probably result in destabilization of the nSec1–syntaxin 1a interaction. Rop and the Ile 236→Ala mutant *Drosophila* syntaxin might still be able to interact with each other, but with reduced affinity.

A model for the role of nSec1 in membrane fusion

The structure of the nSec1–syntaxin 1a complex provides an opportunity to integrate genetic, biochemical and biophysical data into a model that more precisely defines the role of Sec1 proteins in membrane trafficking. The tight association and large contact area between nSec1 and syntaxin 1a, as well as *in vitro* data¹⁴ showing that syntaxin 1a is not capable of forming stable complexes with other proteins when bound to nSec1, imply that a mechanism must exist to dissociate or alter the conformation of the nSec1 so as to permit syntaxin 1a to interact with other proteins. Several lines of evidence indicate that the Rabs or their effector proteins mediate this process. For example, the SLY1–20 mutation uncouples vesicle transport from Rab function³¹ and can bypass the requirement for a Rab in the assembly of the SNARE complex³³. Also, in the yeast trafficking pathway from Golgi to vacuole, the Rab effector Vac1 has been shown to interact with the Sec1 family member Vps45 (ref. 34). Vps45 binds to the yeast syntaxin Pep12 (ref. 35), connecting the

functions of a Rab, a soluble Rab effector and the target-membrane-localized Sec1/syntaxin family members. Although not known to be a Rab effector, *in vitro* data indicate that in *Caenorhabditis elegans* Unc-13 binds to the Sec1 homologue Unc-18 and displaces it from syntaxin³⁶. We suggest that upon recognition of the nSec1–syntaxin 1a complex by a Rab, Rab effector (Fig. 6a), or other factor such as Unc-13, nSec1 changes conformation (Fig. 6b). Based on the limited contact area between nSec1 domain 3a and the rest of the nSec1–syntaxin 1a complex, an outward movement of this domain appears to be a plausible conformational change.

As discussed above, Sec1-like proteins have an active role in trafficking beyond the prevention of syntaxins from binding to partner SNAREs. We hypothesize that when nSec1 changes conformation, it facilitates fusion by providing a platform for SNARE complex assembly. The proposed outward movement of domain 3a would disrupt the favourable environment provided by domain 3a for the syntaxin 1a linker, and when the linker was no longer packed against H3 and Hc, H3 would be more accessible to the SNARE proteins than in the crystal structure observed here. In particular, several syntaxin 1a residues (boxed in Fig. 3b) pack against SNAP-25 in the core complex and form a pair of helices N-terminal to the four-helix bundle³. These residues pack against Habc and the linker region in the nSec1–syntaxin 1a complex. Displacement of the linker region would provide a nucleation site for the SNAP-25 helices (Fig. 6b). As the SNAP-25–syntaxin 1a interaction propagates towards the C terminus of the proteins, residues that interact with VAMP would become accessible, and the VAMP helix would

Table 1 Crystallographic data and refinement statistics

Data collection*									
Data	Wavelength (Å)		Resolution (Å)		% Complete	$R_{\text{sym}}^{\dagger}$	Percentage over $3\sigma(I)$		Average redundancy
MAD	$\lambda_1 = 0.9252$ (remote)		30.0–3.00		98.2 (96.4)	0.062 (0.271)	79.9 (50.3)		4.5 (4.1)
	$\lambda_2 = 0.9795$ (peak)		30.0–3.00		98.4 (97.3)	0.061 (0.276)	80.1 (48.5)		4.5 (4.1)
	$\lambda_3 = 0.9798$ (edge)		30.0–3.00		98.3 (96.7)	0.062 (0.275)	80.2 (48.8)		4.5 (4.1)
Mercuric acetate	1.0006		30.0–3.00		97.3 (98.7)	0.060 (0.272)	76.5 (42.5)		4.0 (4.0)
Uranyl acetate (native of mercuric acetate)	1.0006		30.0–3.10		98.3 (98.7)	0.065 (0.305)	78.2 (48.4)		4.5 (4.4)
Refinement	0.9798		30.0–2.6		98.9 (98.7)	0.058 (0.312)	75.8 (49.9)		5.6 (5.4)
Phasing									
Anomalous diffraction ratios‡									
Wavelength	λ_1		λ_2		λ_3		+ Friedel mate		–Friedel mate
λ_1	0.058		0.058		0.072		1.12		1.47
λ_2			0.070		0.048		0.39		1.04
λ_3					0.056		Reference		0.75
Resolution (Å)	37.11–5.97	5.97–4.74	4.74–4.15	4.15–3.77	3.77–3.50	3.50–3.29	3.29–3.13	3.13–2.99	Overall
Figure-of-merit	0.85	0.79	0.71	0.64	0.54	0.48	0.40	0.32	0.60
Refinement									
R values and temperature factors					Model geometry				
No. of reflections	working set	26,805				Bond length r.m.s.d. from ideal		0.007 Å	
	test set	2,333				Bond angle r.m.s.d. from ideal		1.6°	
$R_{\text{cryst}}^{\text{¶}}$		0.240				Ramachandran plot:#			
$R_{\text{free}}^{\text{¶}}$		0.295				Percentage in most favoured regions			
Average B (Å)	protein	61.6				Percentage in additional allowed regions			
	solvent	32.1				Percentage in generously allowed regions			
Main-chain bond-related B r.m.s.d.		1.6				Anisotropic temperature factor (Å ²):			
Main-chain angle-related B r.m.s.d.		2.3				$B_{11} = B_{22} = 9.0 B_{33} = -18.0$			
Side-chain bond-related B r.m.s.d.		2.7							
Side-chain angle-related B r.m.s.d.		3.6							

R.m.s.d., r.m.s. deviation.

* Values in parentheses are for the highest resolution shell: 3.11–3.00 Å for the MAD phasing data sets, 2.69–2.60 Å for the native data set, 3.11–3.00 Å for the mercuric acetate derivative data set, and 3.21–3.10 Å for the uranyl acetate data set.

† $R_{\text{sym}} = \sum_i \sum_j |I_i(h) - \langle I(h) \rangle| / \sum_i \sum_j I_i(h)$, where $I_i(h)$ is the i th measurement of reflection h , and $\langle I(h) \rangle$ is the weighted mean of all measurements of h . Bijvoet measurements were treated as independent reflections for the MAD phasing data sets.

‡ Anomalous diffraction ratios = $(\Delta F) / \langle |F| \rangle$, where (ΔF) is the r.m.s. Bijvoet difference at a single wavelength (diagonal elements) or the r.m.s. dispersive difference between two wavelengths (off diagonal elements).

§ Phasing power = $\langle |F_H| \rangle / E$, where $\langle |F_H| \rangle$ is the r.m.s. structure factor amplitude for anomalous scatterers and E is the estimated lack-of-closure error. Phasing power is listed for each lack-of-closure expression between the reference data set (+Friedel mate at λ_3) and the + or –Friedel set at each wavelength. Phasing powers were calculated by using all data between 30.0 and 3.0 Å.

¶ The test set comprises a randomly selected subset of the data (7.2%) that was not included in the refinement of the model. The working set contains the remaining reflections from the data set.

¶ $R = \sum_i |F_{\text{obs}}(h_i) - F_{\text{calc}}(h_i)| / \sum_i |F_{\text{obs}}(h_i)|$. R_{cryst} and R_{free} were calculated from the working and test reflection sets, respectively.

As defined in PROCHECK⁴⁶.

nucleate (Fig. 6c). This ordered association, with SNAP-25 preceding the binding of VAMP to syntaxin, is consistent with solution observations of syntaxin–SNAP-25 association in the absence of VAMP^{26,37}. Propagation of the four-helix bundle towards the C terminus could be extremely rapid, and would pull the vesicle and target membrane bilayers into closer proximity (Fig. 6d). It is unclear when nSec1 would be released from the C-terminal portion of H3, as it is not known whether completion of the core complex structure occurs before, coincident with, or after bilayer fusion, but it is conceivable that while nSec1 is associated with this portion of H3 it could assist in promoting membrane fusion³⁸. Furthermore, if nSec1 remained associated with syntaxin through the H3 or Habc regions during fusion, it would remain close to the SNARE complex after fusion and be ideally positioned to rebind syntaxin once the latter had been dissociated from the core complex by NSF. Such sequestering of syntaxin might prevent reassociation of the core complex and ensure the efficiency of NSF and recycling of SNARE proteins.

The notion that nSec1 has an active role in membrane fusion allows us to rationalize the phenotypes of the *Drosophila* Rop mutants described above. According to our model, mutants that display a decrease in neurotransmitter release either cannot bind to syntaxin or, once bound, are incapable of promoting SNARE complex assembly. In contrast, mutants that display an increase in neurotransmitter release might either have a compromised coupling to effector molecules or have fewer contacts with the H3 region of syntaxin such that H3 is more accessible to SNARE proteins. In either case, these mutant proteins would allow a higher rate of SNARE association.

Although various Sec1, syntaxin, VAMP and Rab family members are associated with particular vesicle-trafficking pathways, syntaxin- and VAMP-like proteins associate promiscuously into SNARE complexes *in vitro*. Syntaxin–VAMP pairing in the SNARE core complex is therefore not likely to be the sole determinant of vesicle targeting specificity^{39,40}. Our proposed model allows for specificity in the interactions between particular Sec1–syntaxin complexes and effector molecules. In this way, the Rab–Sec1–Syntaxin system might provide vesicle–target recognition specificity, whereas syntaxin, in conjunction with Sec1, VAMP and SNAP-25 proteins, might be directly involved in the fusion event. □

Methods

Protein expression and purification

Full-length rat nSec1 and the full cytoplasmic domain of rat syntaxin 1a were cloned into the vector pQE9 (Qiagen), using methods described previously^{20,25}. Each protein contains a polyhistidine tag at its N terminus. BL21 *Escherichia coli* cells were transformed with the nSec1-containing vector, and JM109 *E. coli* cells were transformed with the syntaxin 1a vector. Cells were grown in Luria–Bertani medium at 37 °C for ~2 h in the presence of 100 µM ampicillin, and then induced with 1 mM isopropyl thiogalactoside. Cells were harvested by centrifugation, and lysed in a french press. After cell debris had been pelleted, the lysates containing the two different proteins were mixed and the protein complex was affinity-purified on Ni²⁺-NTA beads (Qiagen). The complex was eluted from the affinity column in 400 mM imidazole pH 7.9 and 5 mM 2-mercaptoethanol (2-Me), and the eluate was applied to a Superdex S200 HR26/60 (Pharmacia) column at 4 °C (150 mM NaCl, 5 mM 2-Me, 15 mM Tris-HCl pH 8.2). Fractions containing nSec1–syntaxin 1a complex (~90% pure as determined by SDS–PAGE; data not shown) were concentrated to ~15 mg ml⁻¹ using Centrprep concentrators with a relative molecular weight cut-off of 30,000 (Amicon). Selenomethionine-substituted nSec1 protein was prepared by transforming the *E. coli* B834λ(DE3) methionine auxotroph cells (Novagen) with the pQE9 nSec1 vector, and growing the cells in selenomethionine-containing defined medium⁴¹. Purification of the selenomethionyl-nSec1–syntaxin 1a complex proceeded as described above.

Crystallization

Crystals of the nSec1–syntaxin 1a complex were obtained by vapour diffusion at 21 °C. Almost all of the crystals were extremely thin and exhibited limited diffraction quality, but the use of flat-bottom sitting drop bridges significantly improved the crystals. Purified complex in a buffer containing 150 mM NaCl, 10 mM 2-Me and 15 mM Tris-HCl pH 8.2 was mixed with an equal volume of reservoir solution containing 17% (v/v) PEG 400, 100 mM sodium acetate pH 5.3, 10 mM 2-Me, 10 mM EDTA and 225 mM ammonium

acetate. Crystals appeared within 1 week and grew to an average size of 0.25 × 0.25 × 0.03 mm³ within 2 weeks. The crystals were harvested no later than 2 weeks after they first appeared, and were equilibrated in a solution of 15% (v/v) PEG 400, 100 mM Tris-HCl pH 7.5 and 2 mM 2-Me. This pH adaption procedure produced a 4 Å unit cell change along the *c* axis, so both native and selenomethionine-containing crystals were treated using the same procedure in an attempt to measure isomorphous data sets. Crystals were cryoprotected in Paratone-N (Exxon) and flash frozen at 100 K. The crystals were typically mosaic (~1°) and showed variable diffraction quality. Several hundred crystals were screened to obtain diffraction data suitable for structure determination. The crystals belong to space group P4₂2₁, with unit cell dimensions *a* = *b* = 157.5 Å, *c* = 80.5 Å. The asymmetric unit contains one copy of the nSec1–syntaxin 1a complex, corresponding to a solvent content of 55%. Analysis of washed crystals reveals a mixture of intact and cleaved nSec1. Because the C-terminal portion of the sequence forms an integral part of domain 2 (see text), it is likely that the entire nSec1 polypeptide is present in all copies of the molecule in the crystal, but that the disordered loop is cleaved in some molecules.

Data collection and phasing

All diffraction data used in the determination of the structure were measured at Stanford Synchrotron Radiation Laboratory (SSRL) beamline 9-2 using a Quantum 4 CCD detector (Area Detector Systems Corporation). All data were measured at 100 K, and processed and scaled with Denzo/Scalepack⁴². Radiation damage to frozen crystals limited MAD data collection to three wavelengths, which were systematically measured in 25° wedges with inverse beam geometry. A mercuric acetate derivative data set and two native data sets were also measured. The native data were obtained from crystals that had been soaked in heavy atoms but did not show metal binding as determined by difference Patterson techniques (uranyl acetate for the crystal that served as the native data for the mercuric acetate derivative, and K₂O₈Cl₆ for the 2.6 Å data that were used to refine the final model). Data collection statistics are presented in Table 1.

The mercuric acetate derivative contained a single site, and single isomorphous replacement with anomalous scattering (SIRAS) phases were calculated with the program SOLVE⁴³. Phases from this derivative were used to locate ten selenium sites from the MAD data by difference Fourier methods. Seven additional selenium sites were located by difference Fourier techniques by using phases from the ten-selenium-site model. The 17 (out of a possible 20)-site selenium model was used to calculate initial phases with the Phillips–Hodgson method and maximum-likelihood refinement as implemented in the program CNS⁴⁴. The edge-wavelength (λ₃)-positive Friedel data were used as the reference. The phases were improved by using solvent flipping and histogram matching in CNS, and much of the final model was readily identifiable in the resulting electron density map. An electron density map calculated with combined MAD and SIRAS phases was of only slightly better quality than the MAD map, and was not used for model building.

Model refinement

Model building was done with the program O⁴⁵, and the initial model contained ~75% of the backbone of the two polypeptide chains with much of the sequence assigned. Refinement was performed against all data with $|F| > 0$, with the program CNS⁴⁴. Before refinement, a random subset (~7%) of the data was removed and used as a test set for cross-validation. The model was initially refined by simulated annealing with an MLHL target in CNS against the selenium edge data, with the density-modified phases used as prior phase information. The model was subsequently refined against the 2.6 Å native data set. Subsequent rounds of model building with σ_A -weighted $2F_o - F_c$ maps, positional minimization and individual thermal factor refinement were then performed to produce the final model. Individual thermal factor refinement served to decrease R_{cryst} and R_{free} by ~3.5%. Water molecules were placed in positive-difference Fourier peaks that were greater than 3σ and exhibited a sensible chemical environment. Coordinates and structure factors have been deposited with the Protein Data Bank (accession code 1DN1).

Received 14 December 1999; accepted 29 January 2000.

- Palade, G. Intracellular aspects of the process of protein secretion. *Science* **189**, 347–358 (1975).
- Südhof, T. C. The synaptic vesicle cycle: a cascade of protein-protein interactions. *Nature* **375**, 645–653 (1995).
- Sutton, R. B. *et al.* Crystal structure of a SNARE complex involved in synaptic exocytosis at 2.4 Å resolution. *Nature* **395**, 347–353 (1998).
- Poirier, M. A. *et al.* The synaptic SNARE complex is a parallel four-stranded helical bundle. *Nature Struct. Biol.* **5**, 765–769 (1998).
- Hanson, P. I. *et al.* Structure and conformational changes in NSF and its membrane receptor complexes visualized by quick-freeze/deep-etch electron microscopy. *Cell* **90**, 523–535 (1997).
- Weber, T. *et al.* SNAREpins: minimal machinery for membrane fusion. *Cell* **92**, 759–772 (1998).
- Chen, Y. A. *et al.* The SNARE complex formation is triggered by Ca²⁺ and drives membrane fusion. *Cell* **97**, 165–174 (1999).
- Mayer, A., Wickner, W. & Haas, A. Sec18p (NSF)-driven release of Sec17p (α-SNAP) can precede docking and fusion of yeast vacuoles. *Cell* **85**, 83–94 (1996).
- Söllner, T. *et al.* A protein assembly–disassembly pathway in vitro that may correspond to sequential steps of synaptic vesicle docking, activation, and fusion. *Cell* **75**, 409–418 (1993).
- Pevsner, J., Hsu, S.-C. & Scheller, R. H. n-Sec1: a neural-specific syntaxin-binding protein. *Proc. Natl Acad. Sci. USA* **91**, 1445–1449 (1994).
- Garcia, E. P. *et al.* A rat brain Sec1 homologue related to Rop and UNC18 interacts with syntaxin. *Proc. Natl Acad. Sci. USA* **91**, 2003–2007 (1994).
- Hata, Y., Slaughter, C. A. & Südhof, T. C. Synaptic vesicle fusion complex contains unc-18 homologue bound to syntaxin. *Nature* **366**, 347–351 (1993).
- Pevsner, J. *et al.* Specificity and regulation of a synaptic vesicle docking complex. *Neuron* **13**, 353–361 (1994).
- Yang, B. *et al.* nSec1 binds a closed conformation of syntaxin 1a. *J. Cell Biol.* **148**, 247–252 (2000).

15. Schulze, K. L. *et al.* rop, a Drosophila homolog of yeast Sec1 and vertebrate n-Sec1/Munc-18 proteins, is a negative regulator of neurotransmitter release in vivo. *Neuron* **13**, 1099–1108 (1994).
16. Schekman, R. Genetic and biochemical analysis of vesicular traffic in yeast. *Curr. Opin. Cell Biol.* **4**, 587–592 (1992).
17. Verhage, M. *et al.* Synaptic assembly of the brain in the absence of neurotransmitter secretion. *Science* **287**, 864–869 (2000).
18. Rowe, J. *et al.* Blockade of membrane transport and disassembly of the Golgi complex by expression of syntaxin 1a in neurosecretion incompetent cells: prevention by rbSec1. *J. Cell Sci.* **112**, 1865–1877 (1999).
19. Pfeffer, S. Transport-vesicle targeting: tethers before SNAREs. *Nature Cell Biol.* **1**, 17–22 (1999).
20. Kee, Y. *et al.* Distinct domains of syntaxin are required for synaptic vesicle fusion complex formation and dissociation. *Neuron* **14**, 991–998 (1995).
21. Fernandez, I. *et al.* Three-dimensional structure of an evolutionarily conserved N-terminal domain of syntaxin1a. *Cell* **94**, 841–849 (1998).
22. Dulubova, I. *et al.* A conformational switch in syntaxin1a during exocytosis: role of munc18c. *EMBO J.* **18**, 4372–4382 (1999).
23. Fiebig, K. M. *et al.* Folding intermediates of SNARE complex assembly. *Nature Struct. Biol.* **6**, 117–123 (1998).
24. Tellam, J. T. *et al.* Characterization of Munc-18c and syntaxin-4 in 3T3-L1 adipocytes. *J. Biol. Chem.* **272**, 6179–6186 (1997).
25. Calakos, N. *et al.* Protein–protein interactions contributing to the specificity of intracellular vesicular trafficking. *Science* **263**, 1146–1149 (1994).
26. Nicholson, K. L. *et al.* Regulation of SNARE complex assembly by an N-terminal domain of the t-SNARE Sso1p. *Nature Struct. Biol.* **5**, 793–802 (1998).
27. Fujita, Y. *et al.* Phosphorylation of Munc-18/n-Sec1/rbSec1 by protein kinase C. *J. Biol. Chem.* **271**, 7265–7268 (1996).
28. Fletcher, A. I. *et al.* Regulation of exocytosis by cyclin-dependent kinase 5 via phosphorylation of munc18. *J. Biol. Chem.* **274**, 4027–4035 (1999).
29. Harrison, S. D. *et al.* Mutations in the *Drosophila* Rop gene suggest a function in general secretion and synaptic transmission. *Neuron* **13**, 555–566 (1994).
30. Wu, M. N. *et al.* ROP, the *Drosophila* Sec1 homolog, interacts with syntaxin and regulates neurotransmitter release in a dosage-dependent manner. *EMBO J.* **17**, 127–139 (1998).
31. Dascher, C. *et al.* Identification and structure of four yeast genes (*SLY*) that are able to suppress the functional loss of *YPT1*, a member of the RAS superfamily. *Mol. Cell. Biol.* **11**, 872–885 (1991).
32. Wu, M. N. *et al.* Syntaxin1a interacts with multiple exocytic proteins to regulate neurotransmitter release in vivo. *Neuron* **23**, 593–605 (1999).
33. Lupashin, V. V. & Waters, M. G. t-SNARE activation through transient interaction with a Rab-like guanosine triphosphatase. *Science* **276**, 1255–1258 (1997).
34. Tall, G. G. *et al.* The phosphatidylinositol-3-phosphate binding protein Vac1p interacts with a Rab GTPase and a Sec1p homologue to facilitate vesicle-mediated protein sorting. *Mol. Biol. Cell* **10**, 1873–1889 (1999).
35. Webb, G. C. *et al.* Genetic interactions between a pep7 mutation and the PEP12 and VPS45 genes: evidence for a novel SNARE component in transport between the *Saccharomyces cerevisiae* Golgi complex and the endosome. *Genetics* **147**, 467–478 (1997).
36. Sassa, T. *et al.* Regulation of the UNC-18–*Caenorhabditis elegans* syntaxin complex by UNC-13. *J. Neurosci.* **19**, 4772–4777 (1999).
37. Fasshauer, D. *et al.* A structural change occurs upon binding of syntaxin to SNAP-25. *J. Biol. Chem.* **272**, 4582–4590 (1997).
38. Carr, C. M. *et al.* Sec1p binds to SNARE complexes and concentrates at sites of secretion. *J. Cell Biol.* **146**, 333–344 (1999).
39. Fasshauer, D. *et al.* Mixed and non-cognate SNARE complexes. Characterization of assembly and biophysical properties. *J. Biol. Chem.* **274**, 15440–15446 (1999).
40. Yang, B. *et al.* SNARE interactions are not selective. Implications for membrane fusion specificity. *J. Biol. Chem.* **274**, 5649–5653 (1999).
41. Leahy, D. J. *et al.* Crystallization of a fragment of human fibronectin: introduction of methionine by site-directed mutagenesis to allow phasing via selenomethionine. *Proteins* **19**, 48–54 (1994).
42. Otwinowski, Z. & Minor, W. Processing of x-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).
43. Terwilliger, T. C. SOLVE: An automated crystallographic structure solution program for MIR and MAD. Edition 1.15 (www.solve.lanl.gov, Los Alamos National Laboratory, 1997).
44. Brünger, A. T. *et al.* Crystallography and NMR System (CNS): a new software system for macromolecular structure determination. *Acta Cryst. D* **54**, 905–921 (1998).
45. Jones, T. A. *et al.* Improved methods for the building of protein models in electron density maps and the location of errors in these models. *Acta Cryst.* **A47**, 110–119 (1991).
46. Laskowski, R. A. *et al.* PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* **26**, 283–291 (1993).
47. Kraulis, P. J. MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Cryst.* **24**, 946–950 (1991).
48. Bock, J. B. *et al.* Syntaxin 6 functions in trans-Golgi network vesicle trafficking. *Mol. Biol. Cell* **8**, 1261–1271 (1997).
49. Nicholls, A. GRASP: *Graphical Representation and Analysis of Surface Properties* (Columbia Univ., New York, 1992).

Acknowledgements

We thank H. Bellamy, P. Kuhn, A. Cohen and M. Soltis for SSRL for beamline support; A. Kolatkar, J. Wedekind and members of the Weis laboratory for assistance with data collection and discussions; K. Ervin, R. Hollomon and L. Cai for technical support; E. Hughson for providing Habc coordinates before publication; K. Harlos for providing flat bottom micro-bridges; and A. Brünger, A. May and S. Scales for comments on the manuscript. This work is based on research conducted at SSRL, which is funded by the Department of Energy, Office of Basic Energy Sciences. The Biotechnology Program is supported by the National Institutes of Health, National Center for Research Resources, Biomedical Technology Program and the Department of Energy, Office of Biological and Environmental Research. K.M.S.M. was supported by a Molecular Biophysics training grant from the NIH. This work was supported by grants from the National Institutes of Mental Health to R.H.S. and W.I.W., and also the Pew Scholars Program in the Biomedical Sciences and a Stanford University/Howard Hughes Medical Institute Junior Faculty Award to W.I.W.

Correspondence and requests for materials should be addressed to W.I.W. (e-mail: bill.weis@stanford.edu).

Discovery of a new population of high-energy γ -ray sources in the Milky Way

N. Gehrels, D. J. Macomb, D. L. Bertsch, D. J. Thompson & R. C. Hartman

Laboratory for High Energy Astrophysics, NASA's Goddard Space Flight Centre, Greenbelt, Maryland 20771, USA

One of the great mysteries of the high-energy γ -ray sky is the group of ~ 170 unidentified point sources^{1,2} found along the Galactic plane. They are more numerous than all other high-energy γ -ray sources combined and, despite 20 years of effort, no clear counterparts have been found at other wavelengths. Here we report a new population of such objects. A cluster of ~ 20 faint sources appears north of the Galactic Centre, which is part of a broader class of faint objects at mid-latitudes. In addition, we show in a model-independent way that the mid-latitude sources are distinct from the population of bright unidentified sources along the Galactic plane. The distribution on the sky indicates that the faint mid-latitude sources are associated with the Gould belt^{3,4} of massive stars and gas clouds at ~ 600 light years distance, as has been previously suggested⁵.

We use data from the third source catalogue² of the Energetic Gamma Ray Experiment Telescope (EGRET) on board the Compton Gamma Ray Observatory (CGRO). The catalogue includes observations between 22 April 1991 and 3 October 1995, over which time the instrument⁶ observed the entire sky with exposures in the range $(2-7) \times 10^7$ seconds. Point-source locations and fluxes were determined using a maximum-likelihood approach⁷. Diffuse Galactic and extragalactic γ radiation was included in the analysis based on model parameters⁸. The statistical significance level of sources in the catalogue is 4σ (5σ within 10° of the Galactic plane). The EGRET angular resolution is energy dependent with a full-width at half-maximum of approximately 6° at 100 MeV, decreasing at higher energies. This leads to a typical positioning accuracy for individual sources of $\sim 1^\circ$.

In order to focus on Galactic sources, we eliminate from the current analysis any source with high variability typical of flaring active galactic nuclei. A source is classified as steady if the most significant detection in the EGRET catalogue is for a timescale of years (rather than a single observation of a few weeks) and if that particular flux is within 3σ of the flux calculated for the full data set. The distribution of these 120 steady unidentified EGRET sources on the sky is shown in Fig. 1. It appears from this map that there are two distinct classes of sources, bright ones lying along the Galactic plane

and faint ones at mid-latitudes extending up to $\sim 30^\circ$. The cluster of faint sources above the plane near the Galactic Centre have not shown up so clearly in previous catalogues. Other workers^{9,10} have noted and studied the mid-latitude sources, but the quantitative distinction of these objects as a separate class from the low-latitude sources has not been shown. We report here that the mid-latitude sources are indeed a separate population. (See ref. 11 for discussion of high-latitude unidentified sources and ref. 12 for discussion of sources greater than 1 GeV.)

To quantify the distinction between the low- and mid-latitude sources, we plot in Fig. 2 the flux for each unidentified source as a function of its Galactic latitude as well as the distribution of weak (83) and strong (37) sources defined relative to the mean flux. There is a noticeable change at $|b| \approx 5^\circ$ between bright sources at low latitude and weak sources at mid-latitude. This is not at the 10° latitude point where source significance changes from 4σ to 5σ ; the change at 5° appears to be a real and remarkable effect. Before concluding that the populations are different, we first consider the bias against detecting weak sources along the Galactic plane as a result of it having a diffuse background a factor of ten higher than the mid-latitude sky. To analyse this quantitatively, we plot the integral number of sources brighter than a given flux as a function of that flux ($\log N$ - $\log S$ plot) in Fig. 3. The data appear quite different, with $|b| < 5^\circ$ flattening sharply below 4×10^{-7} ph cm⁻² s⁻¹ compared to only a gradual flattening at much lower flux levels ($< 1 \times 10^{-7}$ ph cm⁻² s⁻¹) for $|b| > 5^\circ$. The solid line shown in both panels is the fit to the high-flux portion of the $|b| > 5^\circ$ data (with same slope but different offset for the $|b| < 5^\circ$ data). If the two populations are the same, then the enhanced flattening for $|b| < 5^\circ$ would be caused solely by detection bias and the real underlying shape would be that shown by the solid line. This cannot be the case because the total flux coming from so many weak sources along the plane would be comparable to the total observed diffuse flux from this region ($|b| < 5^\circ$) of $\sim 2 \times 10^{-4}$ ph cm⁻² s⁻¹ ($E > 100$ MeV; ref 8). The diffuse flux is understood to an accuracy of $\sim 10\%$ as being due to emission from cosmic ray interactions with the interstellar medium, so unresolved point sources cannot be contributing at the level needed to keep the $\log N$ - $\log S$ shapes the same. We conclude that low-latitude and mid-latitude source populations have different intensity distributions: the low-latitude sources are dominated by bright sources, whereas the mid-latitude sources are mostly weak.

Spectral information can also be used to compare the two groups. The average photon spectral index for the low-latitude sources is 2.18 ± 0.04 compared to 2.49 ± 0.04 for the mid-latitudes. These are distinctly different, with a T -test giving a probability of $\sim 10^{-6}$ that the two means are from the same parent sample. We have

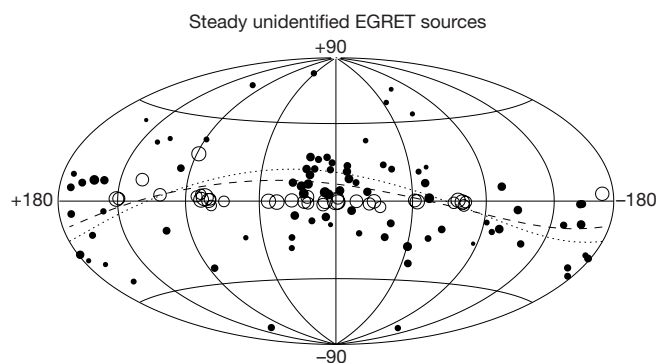


Figure 1 Sky plot of EGRET unidentified sources. The unidentified sources from the third EGRET catalogue² are plotted in an Aitoff projection with the Galactic plane along the central axis. The size of the plotted point for each source is proportional to the flux level of the source. The open (filled) circles are for sources brighter (fainter) than 2.4×10^{-7} ph cm⁻² s⁻¹ ($E > 100$ MeV). Only 'steady' sources are included (see text). The

dotted line shows the approximate centre (O and B star centroid) of the Gould belt. The dashed line is the great circle that best fits the distribution of the EGRET weak sources (filled circles). The two curves are very nearly the same when we take into account that EGRET spent more observing time pointed at the Galactic Centre and anti-centre than at other regions of the sky.

investigated the possibility of a bias from a correlation of gas density with spectral index, but find nothing near the magnitude of the above difference. For example, the difference in index between low-latitude sources near the Galactic Centre direction and those away from it (regions with significantly different gas densities) is less than 0.1. The evidence from spectral analysis of the unidentified sources indicates that the low- and mid-latitude sources are very different in their properties.

We note that the cluster of weak sources north of the Galactic Centre is not caused by an artefact of observational exposure. The EGRET exposure is, in fact, larger at the Galactic Centre than at 20° north or south of the centre by about 50%. The spatial coincidence of the cluster of faint sources above the centre and the diffuse “Galactic halo” found¹³ in the same region bears investigation. We consider two possibilities for the relation between these two components: (1) the point sources are statistical fluctuations in the diffuse flux and not real objects; or (2) the diffuse flux consists of unresolved point sources that are individually below the EGRET sensitivity threshold. To investigate the reality of the point sources, we analysed the point-spread distribution of the individual sources by comparing the reduced χ^2 of fits to the EGRET point-spread function for identified and unidentified sources. The two distributions are consistent, indicating that the unidentified sources are true point sources and not fluctuations in a diffuse flux. To investigate the possibility of the diffuse flux being unresolved point sources, we note that the strength of the diffuse flux¹³ in the $\sim 40^\circ$ -square region of the ‘hot spot’¹³ is $\sim 10^{-6}$ ph cm $^{-2}$ s $^{-1}$. There are 10 EGRET sources in this region with fluxes greater than 10^{-7} ph cm $^{-2}$ s $^{-1}$, so that even a modest extrapolation to weaker sources will clearly give enough flux in the region to explain the diffuse emission. We conclude that the diffuse flux is likely to consist of unresolved point sources from the cluster of faint sources above the Galactic Centre that we discuss here.

Because we have a population of weak, nearby sources at mid-

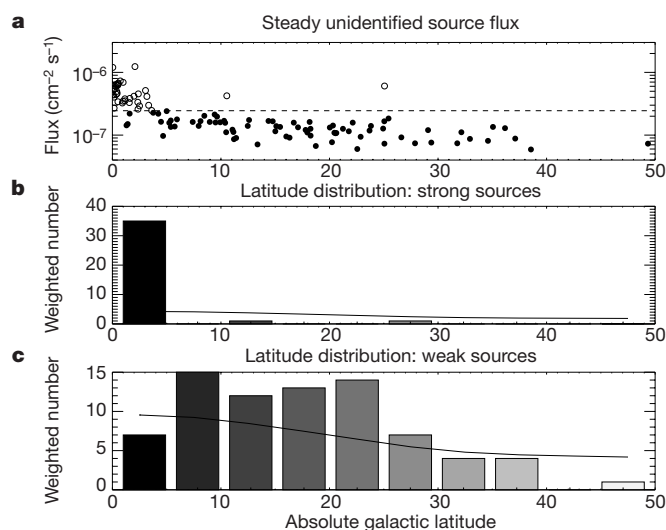


Figure 2 Latitude distribution of the unidentified sources. **a**, The flux levels of the unidentified steady sources from the third EGRET catalogue are shown as functions of absolute Galactic latitude. The open and filled circles have the same definition as in Fig. 1. The flux plotted for each source is the average 1991–1995 flux greater than 100 MeV. The dashed line is the average over the sources shown. The typical flux for sources with $|b| > 5^\circ$ is much lower than for the $|b| < 5^\circ$. **b**, **c**, The distribution in absolute latitude of the unidentified sources with flux above and below the average. Both plots also show the expected number based upon the EGRET exposure as a curved line. The excess of weak sources at mid-latitudes is striking, as is the lack of high-latitude strong sources.

latitudes, we followed the suggestion of other workers (for example, ref. 5) and investigated a Gould belt origin for them. The Gould belt is an asymmetric structure in a great circle on the sky tilted $\sim 20^\circ$ from the Galactic plane. It is comprised of massive¹⁴ and late-type stars⁴, as well as clusters of molecular clouds¹⁵ and expanding interstellar gas¹⁶. Distances to the sources are in the 100–400 pc range. The dotted line on Fig. 1 is the centre of the Gould-belt-associated O and B stars¹⁴. The dashed curve is the best fit to the positions of the EGRET weak sources (that is, filled circles in Fig. 1) for a line of the form $b = A + B\sin(l) + C\cos(l)$, where the fit gives $A = -0.37 \pm 0.11$, $B = 6.62 \pm 0.19$ and $C = 10.02 \pm 0.14$. There is good agreement between the amplitude and phase of the two curves, although the actual belt geometry in terms of ascending node varies with source class^{4,17}. When compared to a fit to constant latitude, an F -test gives a probability of only 2% that the fit improvement due to the two degrees of freedom added by the great circle is due to chance alone.

The dashed data curve in Fig. 1 has a smaller amplitude than the dotted Gould belt curve; this may be caused by the higher EGRET exposure at the Galactic Centre than at 20° north of the centre. This leads to an enhanced density of detected sources at lower Galactic latitude that pulls the data curve down. In addition, the existence of some low-latitude Galactic plane sources in our weak-flux sample will lower the amplitude. The concentration of sources north of the Galactic Centre can be explained by the fact that this is the direction toward the nearest portion of the Gould belt⁴. This interpretation is further supported by looking at the flux distribution of the weak sources. When separated into two samples, those with $60^\circ < l < 240^\circ$ (29 sources) and those with $240^\circ < l < 60^\circ$ (54 sources), the sample around the Galactic Centre has an average flux of 13.5 ± 0.6 whereas the sources towards the anticentre have an average of 11.9 ± 0.7 . The T -test likelihood that these means are from the same population is 5%. This is expected, given the structure of the belt, although the spatial extent of our particular sample is unknown. We

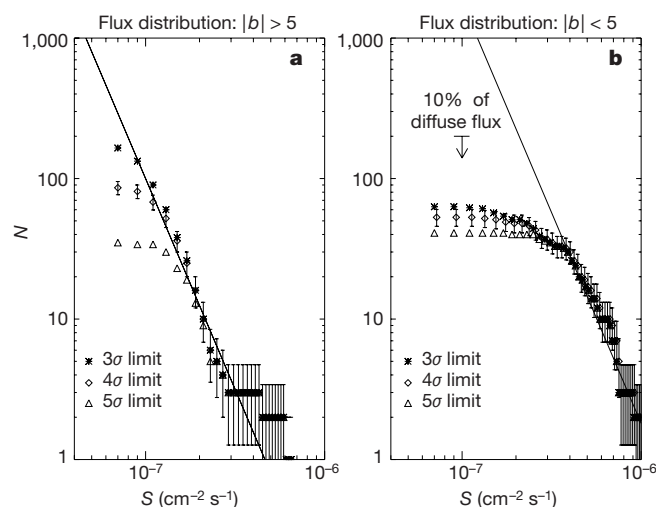


Figure 3 Flux distribution of low- and mid-latitude sources. **a**, **b**, The logN–logS (integral number N versus flux S) distribution of the > 100 MeV EGRET unidentified steady sources is shown between 5° and 30° Galactic latitude (**a**) and along the Galactic plane (**b**). Points are plotted for different significance levels of detection (3σ , 4σ and 5σ), extending to levels below those used for source detection in the catalogue. Error bars are plotted for the 4σ sample only (slightly offset in **b** for clarity); the error bars for the 3σ and 5σ points are similar in size. The solid lines in the two plots represent the best-fit power law (index $= -2.9 \pm 0.7$) to the $|b| > 5^\circ$ data (**a**) and a same-slope power law drawn through the high flux points for the $|b| < 5^\circ$ data (**b**). The upper limit in **b** gives the approximate maximum number of unresolved point sources that could be hidden in the diffuse flux from the region. The $|b| < 5^\circ$ sources have a logN–logS function shape that is clearly different from those at $|b| > 5^\circ$.

conclude that there is evidence, but not proof, that the mid-latitude sources are associated with the Gould belt.

The third EGRET catalogue will be the best data set in high-energy γ -ray astronomy for years to come. (This is because the spark chamber gas in the EGRET instrument is almost completely used up.) With these data we have evidence that the unidentified high-energy γ -ray sources are made up of two separate populations. There must be distinctly different types of objects that make up the two groups. The bright sources are in the Galactic plane and therefore at large (kiloparsec) distances, whereas the weak sources are off the plane and presumably at the 100–400-pc distances of the Gould belt. This implies that the luminosities of the mid-latitude sources are much lower than those of the sources in the Galactic plane. The luminosity difference is of the order of a factor of 3 (for the brightest difference) times 25 (for the square of the distance difference): a total of a factor of ~ 75 .

What kinds of sources are likely to be the counterparts of the mid-latitude objects if they are truly in the Gould belt? Objects that are known to be, or thought likely to be, high-energy γ -ray emitters include molecular clouds¹⁸ (cosmic rays interacting with the enhanced gas concentrations in the clouds to produce γ -ray enhancements in the sky), supernova remnants^{19–22} (cosmic rays accelerated in the supernova explosion interacting with gas in the vicinity to produce γ rays), massive stars^{21,22} (γ rays produced in the outflowing winds from the stars) and pulsars^{10,23,24} (γ rays produced in the particle acceleration regions of the pulsar magnetospheres). Because the Gould belt has an enhanced concentration of massive stars and molecular gas concentrations, we speculate that the γ rays from the mid-latitude sources are produced in these two types of objects. This hypothesis will be tested in 2005 with the launch of the GLAST mission²⁵. Its high sensitivity and improved angular resolution relative to EGRET will allow unique counterparts to be found for the individual unidentified sources. $\square$

Received 8 October 1999; accepted 5 January 2000.

- Swannenburg, B. N. *et al.* Second COS B catalog of high-energy gamma-ray sources. *Astrophys. J.* **243**, L69–L73 (1981).
- Hartman, R. C. *et al.* The Third EGRET catalog of high-energy gamma-ray sources. *Astrophys. J. Suppl.* **123**, 79–202 (1999).
- Gould, B. A. *Brilliant y posición de las estrellas fijas, hasta la séptima magnitud, comprendidas dentro de cien grados del polo austral: con atlas, Uranometria Argentina* (ed. Corni, P. E.) 354 (Buenos Aires, 1879).
- Guillout, P. *et al.* Discovery of a late-type stellar population associated with the Gould Belt. *Astron. Astrophys.* **337**, 113–124 (1998).
- Grenier, I. A. & Perrot, C. Enhanced pulsar and cosmic-ray production in the Gould Belt? *Proc. XXVI Int. Cosmic Ray Conf. Salt Lake City* **3**, 476–479 (1999).
- Thompson, D. J. Calibration of the Energetic Gamma-Ray Experiment Telescope (EGRET) for the Compton Gamma-Ray Observatory. *Astrophys. J. Suppl.* **86**, 629–656 (1993).
- Mattox, J. R. *et al.* The likelihood analysis of EGRET data. *Astrophys. J.* **461**, 396–407 (1996).
- Hunter, S. D. *et al.* EGRET observations of the diffuse gamma-ray emission from the Galactic plane. *Astrophys. J.* **481**, 205–240 (1997).
- Grenier, I. A. Spatial distribution of the unidentified EGRET sources off the Galactic plane. *Adv. Space Res.* **15**, 73–76 (1995).
- Romani, R. W. & Yadigaroglu, I.-A. Gamma-ray pulsars: Emission zones and viewing geometries. *Astrophys. J.* **438**, 314–321 (1995).
- Ozel, M. E. & Thompson, D. J. On the nature of unidentified 2EG sources at high latitudes. *Astrophys. J.* **463**, 105–109 (1996).
- Lamb, R. C. & Macomb, D. J. Point sources of GeV gamma rays. *Astrophys. J.* **488**, 872–880 (1997).
- Dixon, D. D. *et al.* Evidence for a Galactic gamma-ray halo. *New Astron.* **3**, 539–558 (1998).
- Stothers, R. & Frogel, J. A. The local complex of O and B stars: I. Distribution of stars and interstellar dust. *Astron. J.* **79**, 456–471 (1974).
- Taylor, D. K., Dickman, R. L. & Scoville, N. Z. Molecular clouds and the Gould belt. *Astrophys. J.* **315**, 104–121 (1987).
- Lindblad, P. O., Grupe, K., Sandqvist, A. & Schuber, J. On the kinematics of a local component of the interstellar hydrogen gas possibly related to Gould's Belt. *Astron. Astrophys.* **24**, L309–L312 (1973).
- Olano, C. A. On a model of local gas related to Gould's belt. *Astron. Astrophys.* **112**, 195–208 (1982).
- Houston, B. P. & Wolfendale, A. W. Galactic gamma radiation- The contribution from discrete sources. *Astron. Astrophys.* **126**, 22–30 (1983).
- Sturmer, S. J. & Dermer, C. D. Association of unidentified, low latitude EGRET sources with supernova remnants. *Astron. Astrophys.* **293**, L17–L20 (1995).
- Esposito, J. A. *et al.* EGRET observations of radio-bright supernova remnants. *Astrophys. J.* **461**, 820–827 (1996).
- Romero, G. E., Benaglia, P. & Torres, D. F. Unidentified 3EG gamma-ray sources at low Galactic latitudes. *Astron. Astrophys.* **348**, 868–876 (1999).
- Montmerle, T. On gamma-ray sources, supernova remnants, OB associations, and the origin of cosmic rays. *Astrophys. J.* **231**, 95–110 (1979).

- Thompson, D. J. *et al.* EGRET high-energy gamma-ray pulsar studies. I: Young spin-powered pulsars. *Astrophys. J.* **436**, 229–238 (1994).
- Kaaret, P. & Cottam, J. Do the unidentified Galactic EGRET sources lie in star-forming regions? *Astrophys. J.* **462**, L35–L38 (1996).
- Gehrels, N. & Michelson, P. GLAST: the next-generation high energy gamma-ray astronomy mission. *Astroparticle Phys.* **11**, 277–282 (1999).

Acknowledgements

We thank I. Grenier and S. Hunter for discussions.

Correspondence and requests for materials should be addressed to N. G. (e-mail: gehrels@gsfc.nasa.gov).

Trapping an atom with single photons

P. W. H. Pinkse, T. Fischer, P. Maunz & G. Rempe

Max-Planck-Institut für Quantenoptik, Hans-Kopfermann-Strasse 1, 85748 Garching, Germany

The creation of a photon–atom bound state was first envisaged for the case of an atom in a long-lived excited state inside a high-quality microwave cavity^{1,2}. In practice, however, light forces in the microwave domain are insufficient to support an atom against gravity. Although optical photons can provide forces of the required magnitude, atomic decay rates and cavity losses are larger too, and so the atom–cavity system must be continually excited by an external laser^{3,4}. Such an approach also permits continuous observation of the atom's position, by monitoring the light transmitted through the cavity^{5–9}. The dual role of photons in this system distinguishes it from other single-atom experiments such as those using magneto-optical traps^{10–12}, ion traps^{13,14} or a far-off-resonance optical trap¹⁵. Here we report high-finesse optical cavity experiments in which the change in transmission induced by a single slow atom approaching the cavity triggers an external feedback switch which traps the atom in a light field containing about one photon on average. The oscillatory motion of the trapped atom induces oscillations in the transmitted light intensity; we attribute periodic structure in intensity–correlation-function data to ‘long-distance’ flights of the atom between different anti-nodes of the standing-wave in the cavity. The system should facilitate investigations of the dynamics of single quantum objects and may find future applications in quantum information processing.

The interaction of a single two-level atom with a single mode of

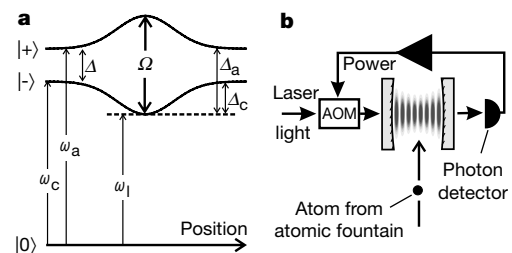


Figure 1 Atom–cavity system. **a**, Ground state and the first excited states of the atom–cavity system as a function of the atom's position in a gaussian mode. The dressed states $|-\rangle$ and $|+\rangle$ are linear combinations of the uncoupled states, with coefficients that depend on Ω . The arrows indicate the frequencies of the atom (ω_a), the cavity (ω_c) and the laser (ω_l). The cavity transmission is high if ω_l is close to one of the dressed states. **b**, Experimental set-up. Atoms are launched by an atomic fountain towards the high-finesse cavity. An acousto-optic modulator (AOM) triggered by the presence of an atom in the cavity increases the light intensity.

the electromagnetic field is well described by the Jaynes–Cummings model¹⁶. The dipole interaction couples the atom and the cavity field, leading to new (dressed) energy eigenstates that are combinations of the bare atom and cavity states. The frequency difference, $\Omega = \sqrt{\Delta^2 + 4g_0^2\psi(\mathbf{r})^2}$, between the two dressed states containing one quantum of excitation is determined by the atom–cavity detuning, $\Delta = \omega_c - \omega_a$, the cavity-field mode function, $\psi(\mathbf{r})$, and the atom–field coupling constant at an antinode, g_0 , where $\psi(\mathbf{r}) = 1$ by definition. Figure 1a displays the dressed energy levels as a function of position for a gaussian mode profile. The pump laser induces transitions between these states with a rate given by the laser intensity, its frequency, ω_b , and the atom’s position, $\mathbf{r}$. The excitation probability determines the cavity transmission and, hence, allows us to observe the atomic motion. Indeed, the transmitted light is a direct measure for the strength of the atom–field coupling. The dipole force on the atom, created by the exchange of a photon between the atom and the cavity field, is given by the negative gradient of the energy of the dressed state, $|+\rangle$ or $|-\rangle$, multiplied by the excitation probability. The energy minimum of state $|-\rangle$, in particular, facilitates three-dimensional atom trapping in the centre of the cavity. In addition to this conservative force, a velocity-dependent force and momentum diffusion induced by spontaneous emission and dipole fluctuations are important. For a weak pump laser, analytical expressions exist for these forces^{17,18}, all confirmed in a recent experiment where the oscillatory motion of atoms channeling through the nodes or antinodes of a standing-wave cavity field containing less than one photon on average was analysed in detail⁴. For larger excitation, more states than just the ground state and the first two excited states must be taken into account. In that case no analytical solutions for the forces are available and the system’s master equation must be solved numerically. Experimentally, cooling and three-dimensional mechanical binding of atoms has not been observed in this system, so far.

We emphasize that the quantum character of the cavity field, that is, its granular nature with associated fluctuations in the number of photons, does not affect the steady-state dipole force and velocity-

dependent force. It does, however, increase the momentum diffusion compared to that in a coherent light field¹⁸. The reason is the back action of the atom on the cavity field, which is not found, for example, in a far-off-resonance dipole trap¹⁵.

The experimental set-up is shown in Fig. 1b. A pulsed atomic fountain launches laser-cooled rubidium (⁸⁵Rb) atoms towards a high-finesse cavity. The flux of atoms is kept so low that at most one atom resides in the cavity at any time. The atoms are optically pumped into the $F = 3$, $m_F = 3$ Zeeman sublevel of the $5^2S_{1/2}$ ground state and the spin orientation is maintained by a small magnetic bias field. The cavity has a finesse of $\sim 4.3 \times 10^5$ and a length of 116 μm , which is actively stabilized, with the exception of the measuring interval of a few milliseconds, during which an atom can be present in the cavity. The cavity field is pumped by a circularly polarized TEM₀₀-mode laser beam, near resonant with the $5^2S_{1/2}$, $F = 3 \leftrightarrow 5^2P_{3/2}$, $F = 4$ transition of the atom at a wavelength of $\lambda = 780$ nm. The intensity and the frequency of this light is controlled by an acousto-optic modulator. The cavity mode function is $\psi(\mathbf{r}) = \cos(2\pi z/\lambda)\exp[-(x^2 + y^2)/w_0^2]$, with waist $w_0 = 29 \mu\text{m}$ and the z -axis horizontal. The light transmitted through the cavity is focused onto a single-photon counting detector with a quantum efficiency of 60%. An additional laser beam resonant with the $5^2S_{1/2}$, $F = 2 \leftrightarrow 5^2P_{3/2}$, $F = 3$ transition of ⁸⁵Rb is injected into the gap between the cavity mirrors. It re-pumps the atom from the uncoupled $5^2S_{1/2}$, $F = 2$ ground state, which would otherwise be populated owing to non-perfect circular polarization of the light in the cavity during the long observation times realized in the experiment. The atom–field coupling constant, $g_0 = 2\pi \times 16$ MHz, atomic dipole decay rate, $\gamma = 2\pi \times 3$ MHz, and cavity field decay rate, $\kappa = 2\pi \times 1.4$ MHz, determine the dimensionless parameters describing the physics of our system: the saturation photon number $\gamma^2/2g_0^2 = 1/57$ and the critical atom number $2\gamma\kappa/g_0^2 = 1/30$. The potential depth for an atom trapped at an antinode of a red-detuned single-photon field is limited to $\hbar g_0$, corresponding to a temperature of 0.8 mK.

To facilitate atom trapping, we first set the atomic fountain to inject very slow atoms (~ 20 cm s⁻¹) into the cavity. Figure 2a displays a transmission signal in the case where we passively observe an atom passing through the cavity. The transit time of this atom amounts to 0.12 ms, consistent with the known entrance velocity. Now we turn on our feedback switch which triggers, at time $t = 0$, when an atom is detected in the cavity, increasing the intensity of the laser light impinging on the cavity within 20 μs . This increase of the potential barrier compensates the kinetic energy of the atom. Figure 2b shows a record of the power transmitted through the cavity in a successful event. The atom that triggered the feedback switch in Fig. 2b remains in the cavity for as long as 0.4 ms. After 1.1 ms, the laser power is switched back to its original value. This allows

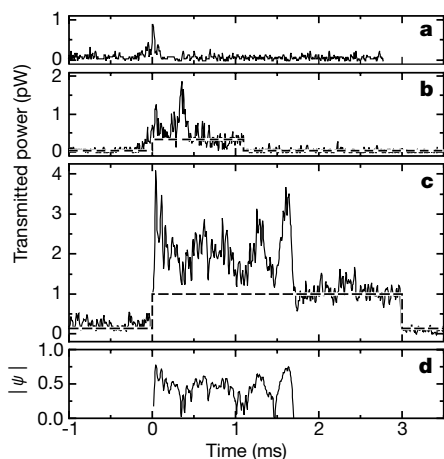


Figure 2 Experimental trajectories. **a**, Transit signal of an atom passing through the high-finesse cavity (solid lines). The detunings are $\Delta_a = \omega_l - \omega_a = -2\pi \times 40$ MHz and $\Delta_c = \omega_l - \omega_c = -2\pi \times 5$ MHz. When pumped with a power of about 150 pW at a detuning of 5 MHz and without the atom, the cavity contains one photon on average, which corresponds to a transmitted power of 0.9 pW. **b**, Same as in **a**, but with feedback. At $t = 0$, the feedback switch triggers an eight-fold increase of the pump power, as is indicated by the dashed line. As a result, the atom causing the trigger remains in the cavity for about 0.4 ms. The pump is switched back to its original value after 1.1 ms. **c**, Same as in **b**, but now with twice the light power and a measurement interval of 3 ms. The atom stays in the cavity for 1.7 ms. The detunings are $(\Delta_a, \Delta_c) = 2\pi \times (-45, -5)$ MHz. **d**, Normalized coupling $|\psi|$ of the atom as inferred from the data in **c**.

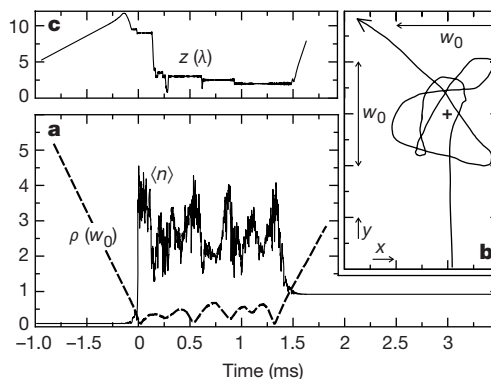


Figure 3 Atomic trajectory calculated with a quantum jump Monte Carlo method. **a**, The solid line shows the mean intracavity photon number and the dashed line indicates the radial ($\rho = \sqrt{x^2 + y^2}$) excursion of the atom. **b**, An axial view. The cross indicates the cavity axis. **c**, The axial (z) motion of the atom.

us to check the length stability of the resonator. The rare events with large resonator drifts are filtered out and discarded.

Demonstration of trapping requires direct evidence for a restoring force, indicated, for example, by an oscillatory motion of the atom. For this purpose, we must examine the transmission signal in more detail. A higher photon flux facilitates this. We therefore slightly increase the laser power in the experiments described below. Figure 2c shows an example where, for a laser intensity twice as large, the atom remains in the cavity for as long as 1.7 ms. This signal is now used to investigate the motion of this particular atom.

As mentioned in the introduction, a transit signal like that in Fig. 2c can directly be mapped onto a time-dependent atom–field coupling, $|\psi\rangle$, is plotted, providing the first information about the trajectory of this particular atom. More physical insight into the atomic motion can be obtained by comparing the experimental data with the results of a numerical simulation. Atomic trajectories are calculated using a quantum-jump Monte Carlo (QJMC) simulation^{19,20}. In this simulation, the atom and the light field are treated quantum mechanically, whereas the motion of the atom is still treated classically. Gravity is neglected, because the optical forces are much larger, even in the weakly confining radial direction. A QJMC simulation is a stochastic solution of the system's master equation, in which the evolution of the quantum state describing the system is computed by combining the coherent evolution with quantum jumps that simulate the spontaneous loss of an excitation out of the system.

Figure 3 shows a result of the QJMC simulation calculated for the parameters of Fig. 2c. Plotted in Fig. 3a are the atom's distance from the cavity axis, $\rho(t)$, and the corresponding mean number of photons in the cavity, $\langle n \rangle$. In this particular simulation, the atom remains trapped for 1.4 ms. During this time interval, the atom oscillates in the radial direction with a period of about 250 μs and an amplitude of about $w_0/2$. The oscillations in the mean photon number clearly correlate with the radial excursions of the atom. From this we conclude that the oscillations observed in the experiment, characterized by the same period and similar intensity changes, are direct evidence for the radial oscillation of the atom. In particular, the transmission is large or small for an atom close to or far away from the cavity axis, respectively.

Figure 3b displays the atomic motion projected onto the xy -plane perpendicular to the cavity axis. The atom enters from below, makes a diffusive motion around the cavity axis, and leaves towards the upper left. We note that the atom's angular momentum changes because of random momentum kicks associated with spontaneous

emission events. In the xy -plane, because of the cylindrical symmetry, only the distance to the cavity axis enters the atom–field coupling. It is therefore not possible to reconstruct the full three-dimensional trajectory of the atom from our experimental data.

We now study motion along the cavity axis. The strong confinement of the atom to regions of length $\lambda/2$ in the axial direction leads to a fast oscillation of the atom with a period of typically a few microseconds. These oscillations, reported in ref. 4, are not visible on the timescale of Fig. 3c. Apart from these oscillations, the atom sometimes escapes from an antinode. This is attributable to dipole fluctuations heating the atom, as mentioned in the introduction. It then hops over a few nodes, is cooled by the friction force and is finally recaptured in another antinode^{17,18}.

Experimental evidence for these 'long-distance' flights comes from the fourth-order intensity-correlation function of the transmitted light, $g^{(4)}(\epsilon, \tau, \epsilon)$, determined from a time-resolved record of the photon-arrival times. This record is used to calculate a second-order intensity-autocorrelation function of events defined by pairs of photons separated by only a short time interval of less than $\epsilon = 150$ ns, discarding isolated photons. As in ref. 4, corrections were made to take into account the finite observation window. Our method of correlating photon pairs enhances intensity fluctuations because the probability of finding a photon pair scales with the square of the light intensity. A typical record from a section, 20 μs in length, of the transit signal from Fig. 2c centred at time $t = 1.51$ ms is displayed in Fig. 4a. These data are now compared with the results of our QJMC simulation, where the atomic trajectory is known. Analysing only a section where the atom moves across several antinodes, as indicated in Fig. 4c, gives the result plotted in Fig. 4b. Both the experimental curve and the simulated curve show a similar periodic modulation. If, in contrast, a section from the simulation with the atom oscillating around an antinode is analysed, hardly any periodic structure is visible. From this we conclude that the periodic modulation in Fig. 4a is direct evidence for the atom moving along the cavity axis.

Based on our detailed understanding of atomic motion, we conclude that atoms with transmission signals like those in Fig. 2b and c are actually trapped in a light field containing one or two photons on average, respectively. When we average over all events where the feedback switch triggered, we find, by fitting an exponential decay curve, an average trapping time of $\tau = 0.25 \pm 0.05$ ms. The same result is found from the QJMC simulation. The trapping time is limited by spontaneous emission kicks causing the atom to escape in the radial direction. This might be prevented by applying feedback-based cooling schemes²¹. One expects quantum aspects in the atomic motion if the atom becomes so cold that a wave description is necessary. It might even be possible to cool the atom into the quantum-mechanical ground state of the optical potential. From the continuous observation of a localized atomic wave packet by means of a quantized light field and the back action resulting from the measurement, new insights into the dynamical behaviour of an individual quantum object can be obtained. Moreover, a single particle trapped in a high-finesse cavity has applications in the rapidly growing field of quantum communication. For example, it should allow the transmission of quantum bits between distant cavities²², to build arbitrary states of the electromagnetic field²³, or to generate a bit stream of single-mode photons^{24,25}. □

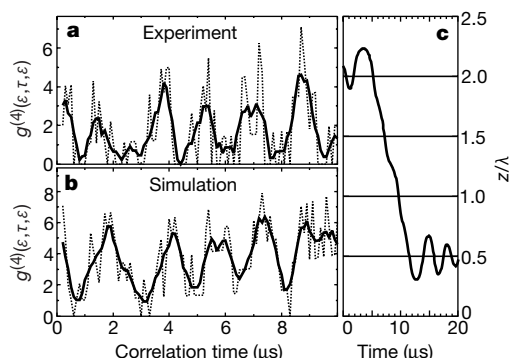


Figure 4 Short-time structure. **a**, **b**, The dashed lines are the fourth-order intensity-autocorrelation function $g^{(4)}(\epsilon, \tau, \epsilon)$. The solid lines are obtained by smoothing. **a** is taken from a section, 20 μs in length, of experimental data centred at 1.51 ms of Fig. 2c; **b** is from an equally long piece of a simulated trajectory, of which the z -coordinate is shown in **c**. The oscillations in **a** and **b** indicate a modulation of the cavity light intensity with a period of 1.8 μs and 2 μs , respectively. The oscillations in **b** are caused by the atom's motion along the z -axis, crossing an antinode (thin lines in **c**) every 2 μs .

Received 8 November 1999; accepted 28 January 2000.

1. Haroche, S., Brune, M. & Raimond, J. M. Trapping atoms by the vacuum field in a cavity. *Europhys. Lett.* **14**, 19–24 (1991).
2. Englert, B.-G., Schwinger, J., Barut, A. O. & Scully, M. O. Reflecting slow atoms from a micromaser field. *Europhys. Lett.* **14**, 25–31 (1991).
3. Hood, C. J., Chapman, M. S., Lynn, T. W. & Kimble, H. J. Real-time cavity QED with single atoms. *Phys. Rev. Lett.* **80**, 4157–4160 (1998).
4. Münstermann, P., Fischer, T., Maunz, P., Pinkse, P. W. H. & Rempe, G. Dynamics of single-atom motion observed in a high-finesse cavity. *Phys. Rev. Lett.* **82**, 3791–3794 (1999).
5. Rempe, G. One atom in an optical cavity: Spatial resolution beyond the standard diffraction limit. *Appl. Phys. B* **60**, 233–237 (1995).

6. Quad, R., Collett, M. & Walls, D. F. Measurement of atomic motion in a standing light field by homodyne detection. *Phys. Rev. Lett.* **74**, 351–354 (1995).
7. Mabuchi, H., Turchette, Q. A., Chapman, M. S. & Kimble, H. J. Real-time detection of individual atoms falling through a high-finesse optical cavity. *Opt. Lett.* **21**, 1393–1395 (1996).
8. Münstermann, P., Fischer, T., Pinkse, P. W. H. & Rempe, G. Single slow atoms from an atomic fountain observed in a high-finesse optical cavity. *Opt. Commun.* **159**, 63–67 (1999).
9. Mabuchi, H., Ye, J. & Kimble, H. J. Full observation of single-atom dynamics in cavity QED. *Appl. Phys. B* **68**, 1095–1108 (1999).
10. Hu, Z. & Kimble, H. J. Observation of a single atom in a magneto-optical trap. *Opt. Lett.* **19**, 1888 (1994).
11. Ruschewitz, F., Bettermann, D., Peng, J. L. & Ertmer, W. Statistical investigations on single trapped neutral atoms. *Europhys. Lett.* **34**, 651–656 (1996).
12. Haubrich, D. *et al.* Observation of individual neutral atoms in magnetic and magneto-optical traps. *Europhys. Lett.* **34**, 663–668 (1996).
13. Neuhauser, W., Hohenstatt, M., Toschek, P. E. & Dehmelt, H. Localized visible Ba⁺ mono-ion oscillator. *Phys. Rev. A* **22**, 1137–1140 (1980).
14. Wineland, D. J. & Itano, W. M. Spectroscopy of a single Mg⁺ ion. *Phys. Lett. A* **82**, 75–78 (1981).
15. Ye, Y., Vernooy, D. W. & Kimble, H. J. Trapping of single atoms in cavity QED. *Phys. Rev. Lett.* **83**, 4987–4990 (1999).
16. Berman, P. R. (ed.) *Cavity Quantum Electrodynamics* (Academic, San Diego, 1994).
17. Horak, P., Hechenblaikner, G., Gheri, K. M., Stecher, H. & Ritsch, H. Cavity-induced atom cooling in the strong coupling regime. *Phys. Rev. Lett.* **79**, 4974–4977 (1997).
18. Hechenblaikner, G., Gangl, M., Horak, P. & Ritsch, H. Cooling an atom in a weakly driven high-Q cavity. *Phys. Rev. A* **58**, 3030–3042 (1998).
19. Carmichael, H. *An Open Systems Approach to Quantum Optics* (Springer, Berlin, 1993).
20. Mölmer, K., Castin, Y. & Dalibard, J. Monte Carlo wave-function method in quantum optics. *J. Opt. Soc. Am. B* **10**, 524–538 (1993).
21. Mancini, S., Vitali, D. & Tombesi, P. Stochastic phase space localization for a single trapped particle. *Phys. Rev. A* (in the press).
22. Cirac, J. I., Zoller, P., Kimble, H. J. & Mabuchi, H. Quantum state transfer and entanglement distribution among distant nodes in a quantum network. *Phys. Rev. Lett.* **78**, 3221–3224 (1997).
23. Law, C. K. & Eberly, J. H. Arbitrary control of a quantum electromagnetic field. *Phys. Rev. Lett.* **76**, 1055–1058 (1996).
24. Law, C. K. & Kimble, H. J. Deterministic generation of a bit-stream of single-photon pulses. *J. Mod. Opt.* **44**, 2067–2074 (1997).
25. Kuhn, A., Hennrich, M., Bondo, T. & Rempe, G. Controlled generation of single photons from a strongly coupled atom-cavity system. *Appl. Phys. B* **69**, 373–377 (1999).

Acknowledgements

We thank P. Münstermann for important contributions to the experiment. The experiments were performed at the University of Konstanz. Funding by the Deutsche Forschungsgemeinschaft, the Optikzentrum Konstanz and the TMR network 'Microlasers and Cavity QED' is gratefully acknowledged.

Correspondence and requests for materials should be addressed to G. R. (e-mail: Gerhard.Rempe@mpq.mpg.de).

An algorithmic benchmark for quantum information processing

E. Knill*, R. Laflamme*, R. Martinez* & C.-H. Tseng†

* Los Alamos National Laboratory, MS B265, Los Alamos, New Mexico 87545, USA

† Department of Nuclear Engineering, MIT, Cambridge, Massachusetts 02139, USA

Quantum information processing offers potentially great advantages over classical information processing, both for efficient algorithms^{1,2} and for secure communication^{3,4}. Therefore, it is important to establish that scalable control of a large number of quantum bits (qubits) can be achieved in practice. There are a rapidly growing number of proposed device technologies^{5–11} for quantum information processing. Of these technologies, those exploiting nuclear magnetic resonance (NMR) have been the first to demonstrate non-trivial quantum algorithms with small numbers of qubits^{12–16}. To compare different physical realizations of quantum information processors, it is necessary to establish benchmark experiments that are independent of the underlying physical system, and that demonstrate reliable and coherent control of a reasonable number of qubits. Here we report an

experimental realization of an algorithmic benchmark using an NMR technique that involves coherent manipulation of seven qubits. Moreover, our experimental procedure can be used as a reliable and efficient method for creating a standard pseudopure state, the first step for implementing traditional quantum algorithms in liquid state NMR systems. The benchmark and the techniques can be adapted for use with other proposed quantum devices.

In NMR experiments the qubits are given by coupled spin-half nuclei in a molecule^{9,10}. A large number of identical molecules are dissolved in a liquid and used as an ensemble of quantum registers. Control is by radio frequency (r.f.) pulses. The initial state is the thermal state and the readout is an ensemble measurement using standard NMR methods. By preparing pseudopure states, it is possible to benchmark quantum algorithms involving up to about ten qubits to determine how reliable the available control methods are. There have been numerous NMR experiments implementing various quantum algorithms^{12–15}. The benchmark proposed and implemented here with seven nuclei requires generating a 'cat state' and then decoding it to the standard initial state. For qubits implemented by spins, the standard initial state has all spins down. The cat state consists of an equal superposition of two states: one with all spins up and the other with all spins down. The cat state is among the most fragile states that are used by quantum computers. A high fidelity realization of our benchmark therefore demonstrates excellent coherent control over the system of qubits.

To simplify the discussion, we use a three-qubit example of the cat-state benchmark. We use deviation density matrices¹⁷ for describing states of the nuclei. This means that states are described by the traceless part of the density matrix up to an overall scale. The thermal equilibrium state of a molecule with one proton (H) and two ¹³C nuclei (C₁ and C₂) at high field in a liquid is given by $\mu_H \sigma_z^{(H)} + \mu_C \sigma_z^{(C_1)} + \mu_C \sigma_z^{(C_2)}$, with μ_H and μ_C the nuclear magnetic moments. The standard Pauli matrices are used as an operator basis, and superscripts on operators refer to the nucleus the operator acts on. The cat-state benchmark for this system begins by eliminating signal from the carbon nuclei to obtain the initial state $\sigma_z^{(H)}$. Next, a sequence of quantum gates¹⁸ is used to achieve the state $\sigma_y^{(H)} \sigma_y^{(C_1)} \sigma_z^{(C_2)}$ (Fig. 1), which is a sum of several coherences¹⁹. In particular, it contains the three coherence $|000\rangle\langle 111| + |111\rangle\langle 000|$, which is the deviation density matrix for the cat state $(|000\rangle + |111\rangle)/\sqrt{2}$ (where

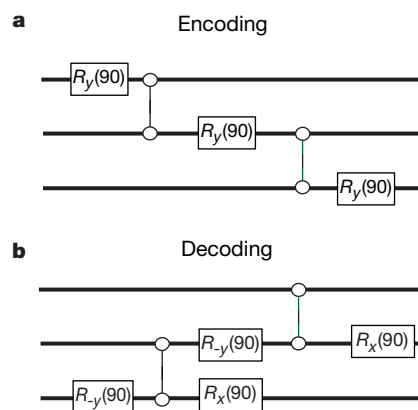


Figure 1 Quantum networks for the cat-state benchmark. **a**, Encoding of the deviation matrix $\sigma_z^{(H)}$ into $\sigma_y \sigma_y \sigma_z$ by using a cascade of one-qubit rotations ($R_y(90)$) and two-qubit operations (vertical bars) given by J -coupling gates. A J -coupling gate is given by the unitary operator $e^{-i\sigma_z \sigma_z \pi/4}$. A three coherence $|000\rangle\langle 111| + |111\rangle\langle 000|$ is contained in the output which can be labelled using a magnetic gradient or phase cycling. **b**, Decoding the coherence to a pseudopure state is accomplished by a similar inverse cascade. The output state is $\sigma_z^{(H)} |00\rangle\langle 00|$. Both networks generalize by extending the cascade to more qubits.

0 and 1 represent the 'down' and 'up' states, respectively). If each qubit is rotated by a phase ϕ around the z-axis, the three coherence rotates by 3ϕ , and all other coherences rotate by 0, ϕ or 2ϕ . This can be used to label the three coherence and eliminate all other components of the state, as explained below. After labelling, the three coherence is decoded to the state $\sigma_z^{(H)}|00\rangle\langle 00|$ (Fig. 1) and observed on the proton. The resulting spectrum is compared to a reference spectrum obtained after applying a 90° r.f. pulse to the proton in the initial, thermal state. The reference spectrum has four peaks corresponding to the states $|00\rangle$, $|01\rangle$, $|10\rangle$ and $|11\rangle$ of the carbon nuclei. The spectrum obtained after decoding the cat state should have a single peak, ideally with the same intensity as that of the corresponding peak in the reference spectrum. If errors in the phase labelling method are negligible, as is the case here, then the ratio of these intensities is a lower bound on the average of the fidelities²⁰ with which the decoding procedure maps the states $|000\rangle \pm |111\rangle$ to the states $|0\rangle \pm |1\rangle|00\rangle$.

The three-qubit cat-state benchmark can be generalized to any number n of qubits by repeating the steps of the cascade in the networks shown in Fig. 1. We implemented the seven-qubit version using fully labelled *trans*-crotonic acid (Fig. 2). The qubits are given by the spin-half component of the methyl protons, the two protons adjacent to the double bond and the four ^{13}C nuclei. Figure 3 shows the experimental reference spectrum of C_1 together with the spectrum obtained after decoding. We chose to observe on the methyl carbon because all the couplings are adequately resolved. Errors can show up as peaks in positions different from the leftmost one. We could not detect such errors above the noise. The fidelity of the experiment was measured to be 0.73 ± 0.02 .

The success of the experiment derives from the use of the following techniques (described in the Methods): (1) An r.f. selection method to greatly reduce the effects of r.f. inhomogeneities. (2) A gradient-based selection method for removing signal from the spin three-half component of the methyl protons in an almost optimal way. (3) The use of abstract reference frames for each nucleus to absorb chemical shift and first-order off-resonance effects in selective pulses. (4) Precomputation of coupling effects during pulses. (5) A pulse-sequence compiler that optimizes delays between pulses for achieving the desired amount of coupling evolution while minimizing unwanted couplings. All of these techniques are scalable in principle.

	M	H ₁	H ₂	C ₁	C ₂	C ₃	C ₄
M	-969.4						
H ₁	6.9	-3560.3					
H ₂	-1.7	15.5	-2938.2				
C ₁	127.5	3.8	6.2	-2327.0			
C ₂	-7.1	156.0	-0.7	41.6	-18599.2		
C ₃	6.6	-1.8	162.9	1.6	69.7	-15412.8	
C ₄	-0.9	6.5	3.3	7.1	1.4	72.4	-21685.1

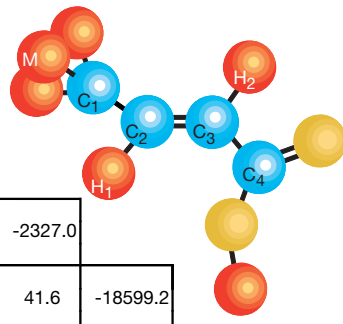


Figure 2 Characteristics of crotonic acid. Molecular structure of *trans*-crotonic acid together with a table of the chemical shifts (on the diagonal) and J -coupling constants (below the diagonal). The chemical shifts are given with respect to reference frequencies of 500.13 MHz (protons) and 125.76 MHz (carbons) on the 500 MHz spectrometer we used. The decoherence times (T_2^*) were greater than 2 s.

The cat-state benchmark has three applications that promise to make it useful for NMR and other quantum technologies. First, the ability to implement the cat-state benchmark with high fidelity is a good indication of what types of tasks can be accomplished in the system at hand. In addition, the decoding algorithm of the cat-state benchmark is similar to necessary subroutines of fault-tolerant error-correction²¹. This is a critical issue for scalable quantum information processing, as robustness requires that each operation has a maximum error below some threshold^{22–24}. Our experiment involved a total of twelve useful two-qubit operations, so the fidelity of 0.73 suggests an error of about 0.023 ($=0.27/12$) per coupling gate. If this degree of control were available in the context of quantum communication, it would be close to the known thresholds²⁵.

Second, the output of the benchmark can be used as a reliable pseudopure state for quantum algorithms using one less ($n - 1$) qubit or as the input to an n -bit error-correction benchmark¹⁴. This requires addressing two problems: the first is to ensure that the labelling method can be used together with a subsequent algorithm, and the second requires that errors accumulated when decoding the n -coherence are eliminated.

The n -coherence can be labelled using gradient methods¹⁹. This has two drawbacks: they are hampered by diffusion and they are not easily generalizable to other devices. One can instead perform $2n + 1$ experiments, where in the k th experiment, the gradient is replaced by explicit pulses that rotate each qubit by a phase $\phi_k = 2\pi k/(2n + 1)$. If o_k is the expectation of the observable measured at the end of the k th experiment, then the value $o = \sum_k o_k e^{-i2\pi kn/(2n+1)}$ is non-zero only for signal originating at an n -coherence. This technique can be applied in any system where it is possible to apply z -rotations reliably. If the phase of applied pulses is

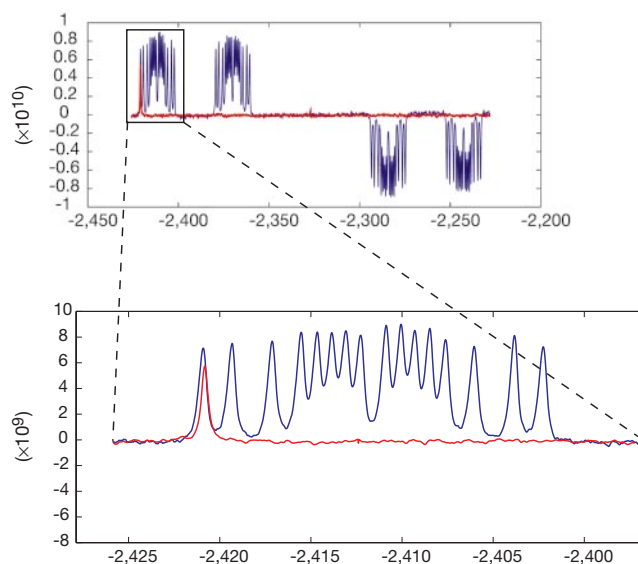


Figure 3 Experimental results: spectra of the pseudopure state (red) and the input state after transfer of polarization to the methyl carbon (blue). The initial steps in the experiments resulting in these spectra were the same, and involved: (1) Signal selection based on radio frequency power. (2) Spin-half selection for the methyl protons. (3) Transfer of polarization to the methyl carbon using one $1/(2J)$ delay. For the pseudopure-state spectrum the next steps were as required for the cat-state benchmark. Both spectra were acquired with 256 scans and are shown at the same scale. The signature of the pseudopure state is that only a single (the leftmost) peak remains. No detectable signal remains in any other peak position. The intensity ratio for the leftmost peak in the pseudopure state and the reference is 0.73 ± 0.02 . This compares well to the relative signal of 0.55 measured for the less fragile states accessed in a five-spins system¹⁶ consisting of four nuclear types. The loss of signal is primarily due to spin relaxation, incomplete refocusing of couplings and intrinsic defects in using selective pulses. Horizontal axes, Hz; vertical axes, intensity (arbitrary units).

highly controllable, as in NMR, one can instead change the reference frame for each qubit, which is equivalent to changing the phase of all subsequent pulses and the observation reference phase by $-\phi_k$. This is essentially a phase-cycling method for selecting the n -coherence²⁶. We have used both the gradient-based and this phase-cycling method with identical results in the crotonic acid system.

The problem of decoding error can in principle be solved by using the maximum coherence directly as the input for a (modified) algorithm. However, it is not possible to obtain a reference signal for the n -coherence without first mapping it to an accessible observable, which can involve a loss of signal. Furthermore, it may be inconvenient to use the n -coherence instead of the more familiar standard pseudopure state. Our experiments show that we can decode the n -coherence to the pseudopure state with no detectable error in the observed spectrum. Errors in unobserved operators can be eliminated efficiently by performing multiple experiments, each with a random phase of 0° or 180° applied to qubits 2, 3, ..., a technique which is a special case of the randomized methods of Knill *et al.*²⁷. The number of experiments that need to be performed depends on the desired level of suppression of possible error signals. N experiments result in suppression by a factor of $O(1/\sqrt{N})$.

The third application of the cat-state benchmark is to demonstrate the ability to reach the maximum coherence with little loss of signal. Previous experiments have generated coherences by exploiting symmetry and effective hamiltonian methods. Very high order coherences can be observed in the solid state²⁶. A maximal coherence of order seven was detected²⁸ by exploiting symmetry. The methods used in these cases yield only a small fraction of the signal that can be achieved by using methods based on quantum networks.

The realization of the cat-state benchmark given here is in an ensemble setting. Most proposals for quantum devices involve individual systems with pure initial states. In these cases the benchmark can be modified by replacing the ensemble measurements by repetition to infer σ_k with sufficiently high signal-to-noise ratio. The preparation step is replaced by a network that directly maps the available initial state to the cat state. We note that any evaluation of a quantum device involves substantial repetition, essentially replacing the ensemble measurement by an ensemble in time. □

Methods

We used a standard 500 MHz NMR spectrometer (DRX-500 Bruker Instruments) with a triple resonance probe to characterize ^{13}C -labelled *trans*-crotonic acid (Fig. 2) at 298K.

To make crotonic acid-1,2,3,4- $^{13}\text{C}_4$ we started from ethyl crotonate-1,2,3,4- $^{13}\text{C}_4$ which was reduced with an excess of sodium borohydride and the alcohol formed was converted to the tosylate in near quantitative yield. The tosylate (unpurified) was then refluxed in ethanol with potassium carbonate to give predominantly the ethyl *trans*-crotonate-1,2,3,4- $^{13}\text{C}_4$ (ratio of 97:3 to *cis*). Crotonic acid-1,2,3,4- $^{13}\text{C}_4$ was obtained in 91% overall yield by saponification of the ester with aqueous sodium hydroxide in water.

Selective r.f. pulses used in the experiment were designed and analysed by simulation on single and pairs of nuclei and represented optimally as a composition of phase shifts, $\sigma_x\sigma_z$ couplings and an ideal 90° or 180° pulse. The simulation is in principle scalable. This permits elimination of most first-order errors attributable to off-resonance and coupling effects without using specialized shapes.

To implement quantum information processing tasks, we translated an ideal quantum network, expressed in terms of 90° rotations and $1/(2J)$ coupling evolutions, to a pulse sequence. After inserting suitable refocusing pulses, a home-built pulse-sequence compiler optimized the delays between pulses to achieve the desired evolution and minimize errors due to incompletely refocused couplings. Coupling evolutions were calculated based on delays and first-order effects during selective pulses. Optimization involved minimizing the sum of the squares of the predicted coupling over- or under-rotation for each pair of nuclei and step of the quantum network. The final pulse sequence used in our experiment required 48 pulses with a signal loss caused by coupling errors (explicitly estimated by the compiler) of 0.15. Phases for each pulse were with respect to phases of abstract reference frames for each nucleus. The phases of the abstract reference frames were precomputed based on chemical shift evolution and first-order side-effects caused by selective pulses.

The effect of pulse imperfections due to r.f. inhomogeneities was reduced by selecting signal based on r.f. power. The method exploits the dependence of the nutation rate on r.f. power and is similar to previously implemented r.f. imaging methods²⁹. Our implementation consisted of the following sequence applied to the methyl protons:

$90^\circ_x - (180^\circ_{-x})^{64} - (180^\circ_{\phi_1} - 180^\circ_{-\phi_1})^{64} - 90^\circ_y$. The phases ϕ_i were modulated by a 64-

point gaussian. The sum of these angles was 22.5° . To remove unwanted signal, each experiment was repeated twice, with the selection sequence in the second experiment using $-\phi_i$ instead of ϕ_i . The signals from the two experiments were subtracted. The signal selected by this sequence is r.f.-homogeneous to within $\pm 2\%$. By calibrating the power, we were able to retain about 25% of the signal compared to an unselected spectrum. This sequence also selects signal from only the methyl protons so that the initial state is $\sigma_z^{(M)}$.

To use the three methyl protons as a single effective spin-half particle, we exploited the fact that, owing to symmetry, the state space of the protons consists of one spin three-half component and two spin-half components. The initial state involves polarization in both spin types. It suffices to eliminate signal from the first while retaining polarization from the second. This was accomplished by using a three-step sequence involving transfer of polarization to the adjacent carbon and terminated by a gradient 'crusher' (details available from the authors). The elimination of signal from the spin three-half states was verified by observing the signal on the adjacent carbon after transfer of the methyl polarization with different coupling delays. Signal from the spin three-half component shows up in the extreme pair of the four main peaks given by the coupling to the methyl protons. We were unable to detect any error signal above the noise.

The remaining steps of the experiment consist of generation and labelling of the n -coherence, and decoding operations to obtain the standard pseudopure state. The sequence is as described, with inserted refocusing pulses and optimized delays. Knowledge of the intended current state of a nucleus was used when that state was $|0\rangle\langle 0|$ or $|1\rangle\langle 1|$, to absorb the effects of couplings to that nucleus in the reference frame³⁰.

Received 30 August 1999; accepted 27 January 2000.

- Simon, D. R. On the power of quantum computation. *SIAM J. Comput.* **26**, 1474–1483 (1997).
- Shor, P. W. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM J. Comput.* **26**, 1484–1509 (1997).
- Wiesner, S. Conjugate coding. *Sigact News* **15**, 78–88 (1983).
- Bennett, C., Bessette, F., Brassard, G., Salvai, L. & Smolin, J. Experimental quantum cryptography. *J. Cryptol.* **5**, 3–28 (1992).
- Cirac, J. & Zoller, P. Quantum computations with cold trapped ions. *Phys. Rev. Lett.* **74**, 4091–4094 (1995).
- Loss, D. & DiVincenzo, D. P. Quantum computation with quantum dots. *Phys. Rev. A* **57**, 120–126 (1997).
- Bocko, M. F., Herr, A. M. & Feldman, M. J. Prospects for quantum coherent computation using superconducting electronics. *IEEE Trans. Appl. Supercond.* **7**, 3638–3641 (1997).
- Shnirman, A., Schön, G. & Hermon, Z. Quantum manipulations of small josephson junctions. *Phys. Rev. Lett.* **79**, 2371–2374 (1997).
- Cory, D. G., Fahmy, A. F. & Havel, T. F. Ensemble quantum computing by nmr-spectroscopy. *Proc. Natl Acad. Sci. USA* **94**, 1634–1639 (1997).
- Gershenfeld, N. A. & Chuang, I. L. Bulk spin resonance quantum computation. *Science* **275**, 350–356 (1997).
- Kane, B. E. A silicon-based nuclear spin quantum computer. *Nature* **393**, 133–137 (1998).
- Jones, J. A., Mosca, M. & Hansen, R. H. Implementation of a quantum search algorithm on a quantum computer. *Nature* **392**, 344–346 (1998).
- Chuang, I. L., Vandersypen, L. M. K., Zhou, X., Leung, D. W. & Lloyd, S. Experimental realization of a quantum algorithm. *Nature* **393**, 143–146 (1998).
- Cory, D. G. *et al.* Experimental quantum error correction. *Phys. Rev. Lett.* **81**, 2152–2155 (1998).
- Nielsen, M. A., Knill, E. & Laflamme, R. Complete quantum teleportation. *Nature* **396**, 52–55 (1998).
- Marx, R., Fahmy, A. F., Myers, J. M., Bermel, W. & Glaser, S. J. Realization of a 5-bit nmr quantum computer using a new molecular architecture. Preprint quant-ph/9905087 at (<http://xxx.lanl.gov>) (1999).
- Sörensen, O. W., Eich, G. W., Levitt, M. H., Bodenhausen, G. & Ernst, R. R. Product operator-formalism for the description of nmr pulse experiments. *Prog. Nucl. Magn. Reson. Spectrosc.* **16** (1983).
- Barenco, A. *et al.* Elementary gates for quantum computation. *Phys. Rev. A* **52**, 3457–3467 (1995).
- Freeman, R. *Spin Choreography* (Oxford Univ. Press, 1998).
- Schumacher, B. Sending entanglement through noisy quantum channels. *Phys. Rev. A* **54**, 2614–2628 (1996).
- Shor, P. W. in *Proc. Symp. Foundations of Computer Science* 56–65 (IEEE, Los Alamitos, California, 1996).
- Aharonov, D. & Ben-Or, M. in *Proc. 29th Ann. ACM Symp. Theory of Computing* 176–188 (Association for Computing Machinery, New York, 1996).
- Kitaev, A. Y. Quantum computations: algorithms and error correction. *Uspekhi Mat. Nauk.* **52**, 53–112 (1997).
- Knill, E., Laflamme, R. & Zurek, W. H. Resilient quantum computation. *Science* **279**, 342–345 (1998).
- Dür, W., Briegel, H.-J., Cirac, J. I. & Zoller, P. Quantum repeaters for based on entanglement purification. *Phys. Rev. A* **59**, 169–181 (1999).
- Emsley, L. & Pines, A. in *Proc. Int. School of Physics, Enrico Fermi Vol. CXXIII*, 123–266 (North-Holland, Amsterdam, 1994).
- Knill, E., Chuang, I. & Laflamme, R. Effective pure states for bulk quantum computation. *Phys. Rev. A* **57**, 3348–3363 (1998).
- Weitekamp, D. P., Garbow, J. R. & Pines, A. Determination of dipole coupling constants using heteronuclear multiple quantum nmr. *J. Chem. Phys.* **77**, 2870–2883 (1982).
- Maffei, P., Elbayed, K., Brondeau, J. & Canet, D. Slice selection in nmr imaging by use of the b1 gradient along the axial direction of a saddle-shaped coil. *J. Magn. Reson.* **95**, 382–386 (1991).
- Vandersypen, L. M. K., Yannoni, C. S., Sherwood, M. H. & Chuang, I. L. Realization of logically labeled effective pure states for bulk quantum computation. *Phys. Rev. Lett.* **83**, 3085–3088 (1999).

Acknowledgements

We thank D. Cory, G. Fernandez, T. Havel, S. Lacelle, D. Lemaster, D. Meyer, C. Unkefer, A. M. Wang and W. H. Zurek for help and discussions.

Correspondence and requests for materials should be sent to E.K. (e-mail: knill@lanl.gov).

Storage of X-ray photons in a crystal resonator

K.-D. Liss*, R. Hock†, M. Gomm†, B. Waibel‡, A. Magerl†, M. Krisch* & R. Tucoulou*

* European Synchrotron Radiation Facility, B.P. 220, F-38043 Grenoble Cedex, France

† Lehrstuhl für Kristallographie und Strukturphysik, D-91054 Erlangen, Germany

‡ MTU (Motoren- und Turbinen-Union) GmbH, D-80991 München, Germany

The temporal structure and high brilliance of the X-ray beams produced by third-generation synchrotrons open up new possibilities in time-dependent diffraction and spectroscopy, where timescales down to the sub-nanosecond regime can now be accessed. These beam properties are such that one can envisage the development of the X-ray equivalent of optical components, such as photon delay lines and resonators, that have proved indispensable in a wide range of experiments—for example, pump-probe and multiple-interaction experiments—and (through shaping the temporal structure and repetition rate of the beams) time-dependent measurements in crystallography, physics, biology and chemistry^{1–3}. Optical resonators, such as those used in lasers, are available at wavelengths from the visible to soft X-rays^{4,5}. Equivalent components for hard X-rays have been



Figure 1 Sketch of the resonator with the two active slices sticking out of the monolithic device. Photons impinging from the left can be trapped due to transmission and reflection probabilities T and R , respectively. They exit the resonator with probabilities TT , $TRRT$, $TRRRRT$, ..., T^2R^{2N} for $0, 1, 2, \dots, N$ back-and-forth bounces at $t_0, t_1, \dots$, being delayed by N times the time of flight inside the resonator of 1.0 ns. The reflection angle is adjusted to exactly 180 degrees and the beam paths of multiply reflected photons are exactly superimposed in space.

discussed for more than thirty years^{4,6,7}, but have yet to be realized. Here we report the storage of hard X-ray photons (energy 15.817 keV) in a crystal resonator formed by two plates of crystalline silicon. The photons are stored for as many as 14 back-and-forth cycles within the resonator, each cycle separated by one nanosecond.

The X-ray resonator consists of a pair of vertical plates cut from a monolithic silicon crystal, separated by 150 mm and with the 111 orientation along their surface normals, as sketched in Fig. 1. The plates are slightly wedge-shaped in order to vary the effective crystal thickness between 50 μm and 500 μm by a horizontal translation perpendicular to the axis of the beam. The experiment was performed at the inelastic scattering beamline ID28 at the European Synchrotron Radiation Facility (ESRF). The radiation from the undulator source was monochromatized using the silicon 888 reflection at a Bragg angle of 89.865 degrees. This provided an X-ray beam of approximately 15.817 keV with an energy resolution of 3.7 meV and a divergence of about 10 mrad. The energy of the incoming photons could be varied through thermal expansion of the lattice spacing by an accurate temperature control of the monochromator, and a typical overall stability of 10 mK during the integration time of one pattern could be achieved. (Using the thermal expansion coefficient $\alpha = 2.56 \times 10^{-6} \text{ K}^{-1}$ for silicon, gives the relative lattice parameter uncertainty $\Delta d/d = \alpha \Delta T = 2.56 \times 10^{-8}$, where T is temperature. This is a much smaller value than the energy resolution $\Delta E/E = 2.3 \times 10^{-7}$ of the incident X-ray beam.) The resonator was aligned such that photons are Bragg reflected exactly back into the axis of the incident beam, and the intensity transmitted through the crystal plates was measured behind the resonator by a fast avalanche diode with a response time matching the intrinsic time structure of the synchrotron X-ray beam.

The Bragg condition for the resonator was determined by energy variation of the incoming photons through the controlled thermal expansion of the monochromator lattice spacing. The result of such a scan is shown in Fig. 2. The exact Bragg condition is fulfilled when the transmitted intensity has a minimum, because most of the photons are then back reflected and do not reach the detector. The observed full-width-half-maximum (FWHM) on the relative energy scale is $\Delta E/E = 7.4 \times 10^{-7}$ and the minimum transmission through the two 292- μm slices is 17% after consideration of normal absorption.

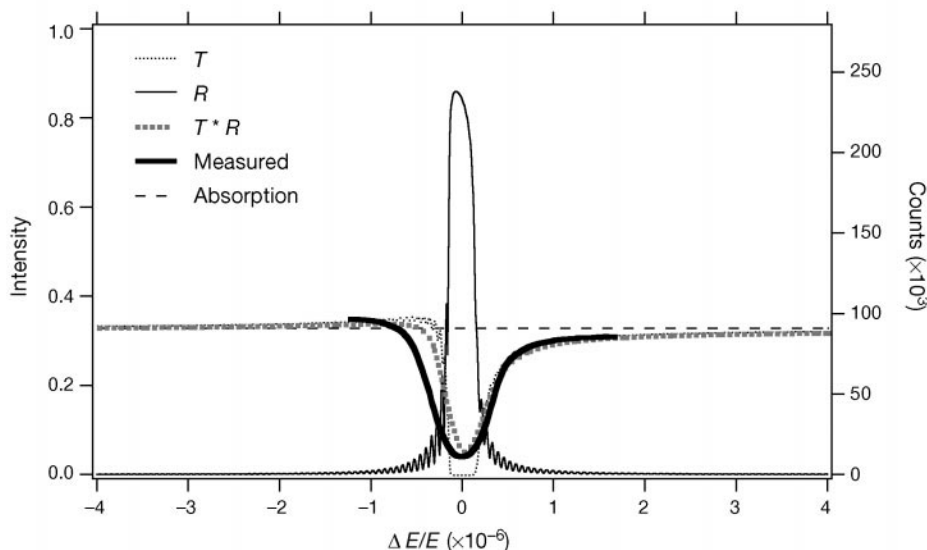


Figure 2 Calculated reflection and transmission curves R and T of a resonator $2 \times 292 \mu\text{m}$ in thickness as a function of relative energy deviation $\Delta E/E$ from the centre of the Bragg reflection. The monochromator delivers a similar curve to R , regardless of the small wiggles stemming from the slice thickness; thus the calculated curve for the simulation of

an energy scan is the convolution T^*R . The experimental data have been adapted to the intensity scale. The dashed line is the absorption far off the Bragg peak; the experimental data is consistent with theory for anomalous absorption, the asymmetry around the Bragg position.

The time dependence of the transmitted beam at the centre of the Bragg reflection, that is, at the dip in Fig. 2, is shown in Fig. 3 for various crystal thicknesses. The data was obtained in the ESRF 16-bunch mode where X-ray pulses of 100-ps duration are separated by 176 ns. The time delay between a registered photon and the synchrotron bunch clock was measured stroboscopically and the events were stored in corresponding channels of a multi-channel analyser³. The first curve shows the time structure of the detector without the resonator in the beam. The intensity maximum corresponds to the undelayed photons at time $t = 0$. The signal has a FWHM of 500 ps, corresponding to the time response of the detector. The shape asymmetry arises from intrinsic capacities⁸ and can be neglected in this experiment. The time patterns with the resonator in the Bragg position differ qualitatively from the first curve and show a series of sequential maxima separated by 1.0 ns forming an exponential decay towards longer times. The maxima correspond to photons trapped within the resonator for 1, 2, 3... N successive reflections from both crystal plates. When the beam impinges onto the first crystal slice, there is a probability for transmission, giving the forward diffracted beam. The same holds

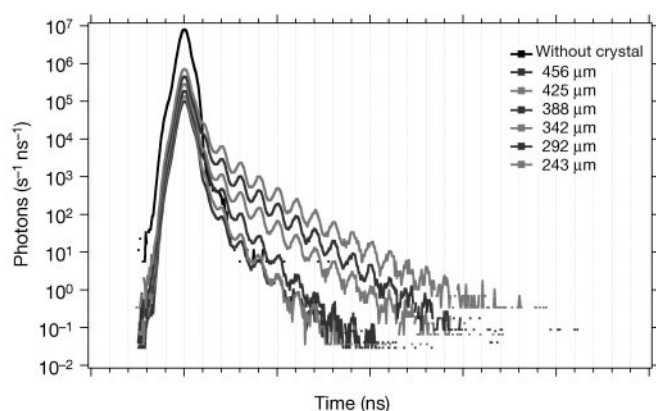


Figure 3 Time pattern of stored photons at the exit of the resonator. Without the crystal in the beam, only a direct bunch of photons is observed at time zero, whereas many bunches appear when the device is in Bragg condition. Each peak at later times corresponds to multiple back-and-forth travels inside the device. The widths of the peaks are dominated by the time resolution of the avalanche detector.

for the second plate and therefore yields a maximum signal at $t = 0$. But there is also a probability for reflection at each slice, permitting part of the beam to go back and forth and thus travel several times the length between the slices, that is, multiples of 30 cm in length or of 1.0 ns in time, resulting in the detected peak separations. We observe up to 14 peaks corresponding to up to 14.0 ns of time delay and even intensity beyond that for the 243- μm -thick crystals. The delayed maxima are less intense for thicker crystal slices because the transmission probability for entering and leaving the resonator is reduced, even in the case of no absorption, and a compromise has to be made between transmission and reflectivity.

The storage capacity of the resonator depends on the ratio of the crystal thickness to the 'Pendellösung period' of the chosen reflection⁹; this ratio defines the minimum thickness that will yield maximum reflectivity. For the silicon 888 reflection it is 150 μm ; a resonator of this plate thickness would reach optimum storage and thus longest delay times. For thicker crystals, fewer photons enter or exit the resonator, as a result of back reflection, and so the total observed intensity decreases. Thinner crystals down to half a Pendellösung period could be useful for a small number of bounces, that is, a small number of reflections, when higher transmission is advantageous at the cost of reduced reflectivity.

The intensity ratios between successive back-and-forth bounces derived from Fig. 3 are plotted in Fig. 4 as a function of the number of bounces. Again, the thinnest crystal gives the highest values. Most of the intensity leaks out of the resonator during the first bounces, mainly as photons with energies in the tails of the reflection curve where the reflectivity is low. Photons exactly fulfilling the Bragg condition have the highest reflectivity and thus survive longest in the cavity. Consequently, the beam becomes first highly monochromatic, reducing the intensity of its wings, and only photons in a very narrow energy range within the intrinsic width of the reflection curve remain in the resonator. The experimental bounce ratio approaches asymptotically the value 0.5, corresponding to a crystal reflectivity of about 70%. The points connected by the continuous line are results from a calculation. The curve shows a faster decrease for the very first bounces followed by a higher asymptotic value of 66%. These discrepancies with theory are consistent with the broader width and deeper dip in the experimental energy scan of Fig. 2 and are probably attributable to lattice strain of the Czochralski grown crystal with high oxygen concentration used in this set-up.

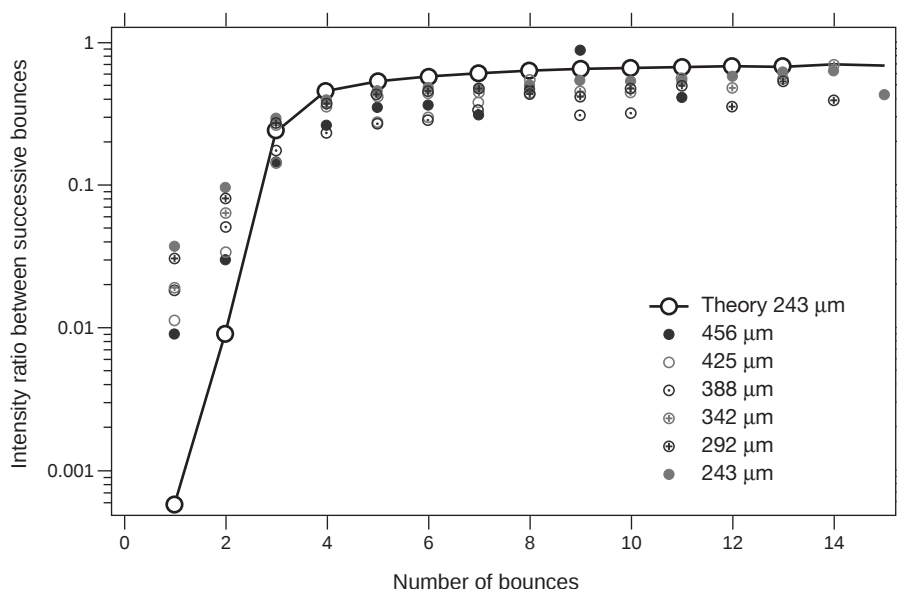


Figure 4 Intensity ratios between neighbouring peaks in the time period of Fig. 3 and calculated values obtained from dynamical theory (joined circles).

In the present set-up, photons are stored by reflections from two independent Bragg mirrors, that is, the photon pulse length is shorter than the resonator length. We envisage the construction of a Fabry–Perot-type interferometer where the incoming beam would interfere coherently with both crystal slabs. This additional interference would lead to a transmission resonance at least ten times sharper than the natural Bragg line width. Such devices necessitate a gap width of a few Pendellösung periods, that is, a few hundred micrometres (refs 10 and 11). With the development of free electron lasers, pulses on the 100-fs scale will become available—matching these distances perfectly. □

Received 7 September 1999; accepted 20 January 2000.

1. Perman, B. *et al.* Energy transduction on the nanosecond time scale: early structural events in a xanthopsin photocycle. *Science* **279**, 1946–1950 (1998).
2. Rose-Petruck, C. *et al.* Picosecond-milliÅngström lattice dynamics measured by ultrafast X-ray diffraction. *Nature* **398**, 310–312 (1999).
3. Liss, K.-D., Magerl, A., Hock, R., Waibel, B. & Remhof, A. The investigation of ultrasonic fields by time resolved X-ray diffraction. *Proc. SPIE* **3451**, 117–127 (1998).
4. Bond, W. L., Duguay, M. A. & Rentzepis, P. M. Proposed resonator for an X-ray laser. *Appl. Phys. Lett.* **10**, 216–218 (1967).
5. Ceglie, N. M., Gaines, D. P., Trebes, J. E., London, R. A. & Stearns, D. G. Time-resolved measurement of double pass amplification of soft X-rays. *Appl. Opt.* **27**, 5022–5025 (1988).
6. Deslattes, R. D. X-ray monochromators and resonators from single crystals. *Appl. Phys. Lett.* **12**, 133–135 (1968).
7. Hart, M. Bragg reflection X-ray optics. *Rep. Prog. Phys.* **34**, 435–490 (1971).
8. Kishimoto, S. An avalanche photodiode detector for X-ray timing measurements. *Nucl. Instrum. Methods Phys. Res. A* **309**, 603–605 (1991).
9. Zachariasen, W. H. *Theory of X-Ray diffraction in Crystals* (Wiley, London, 1945).
10. Steyerl, A. & Steinhauser, K.-A. Proposal of a Fabry–Perot-type interferometer for X-rays. *Z. Phys. B* **34**, 221–227 (1979).
11. Caticha, A. & Caticha-Ellis, S. A Fabry–Perot interferometer for hard X-rays. *Phys. Status Solidi A* **119**, 643–654 (1990).

Acknowledgements

We thank the European Synchrotron Radiation Facility crystal laboratory for the preparation of the resonator.

Correspondence and requests for materials should be addressed to K.-D.L. (e-mail: liss@esrf.fr).

Soft-mode hardening in SrTiO₃ thin films

A. A. Sirenko*, C. Bernhard†, A. Golnik†, Anna M. Clark*, Jianhua Hao*, Weidong Si* & X. X. Xi*

* Department of Physics, The Pennsylvania State University, University Park, Pennsylvania 16802, USA

† Max-Planck-Institut für Festkörperforschung, Heisenbergstraße 1, D-70569 Stuttgart, Germany

Understanding the behaviour of the dielectric constant in ferroelectric thin films remains a challenging problem. These ferroelectric materials have high static dielectric constants, and so are important for their applications in high-storage-density capacitor structures such as dynamic random access memory (DRAM)¹. But the dielectric constant tends to be significantly reduced in thin films, thereby limiting the potential benefit of ferroelectrics for memory devices². Extensive studies have shown that this phenomenon could be caused by a ‘dead layer’ of very low dielectric constant between the ferroelectric film and the electrode^{2,3}. And, although very few direct measurements are in fact available, it has been recognized that the lattice dynamical properties in the thin films should also play a key role in the reduction² of the dielectric constant. Here we report far-infrared ellipsometry and low-frequency dielectric measurements in SrTiO₃ thin films, which

demonstrate that the Lyddane–Sachs–Teller relation between the optical-phonon eigenfrequencies and the dielectric constant is fully maintained, as is the case in the bulk material. This indicates that the dramatic reduction of the dielectric constant is a consequence of a profound change of the lattice dynamical properties, in particular of the reduced softening of its lowest optical-phonon mode. Our results therefore provide a better understanding of the fundamental limitations of the dielectric constant values in ferroelectric thin films.

Phonons are important in the phase transitions in the ferroelectric perovskite titanates SrTiO₃ (STO) and (Ba,Sr)/TiO₃ (BST)^{4,5}. As temperature decreases, the eigenfrequency of the lowest optical mode (the soft mode) falls and approaches zero at a critical temperature T_c where a lattice instability leads to a ferroelectric phase transition⁶. The Lyddane–Sachs–Teller (LST) relation for a crystal with N infrared-active optical modes ($N = 3$ for STO) is:

$$\frac{\epsilon(0)}{\epsilon(\infty)} = \prod_j \frac{\omega_{LOj}^2}{\omega_{TOj}^2}$$

This relates the static dielectric constant, $\epsilon(0)$, and the high-frequency dielectric constant, $\epsilon(\infty)$, to the eigenfrequencies, ω_{LOj} and ω_{TOj} , of the longitudinal (LO) and transverse (TO) optical-phonon modes, respectively. It is generally found that the eigenfrequencies of the higher optical modes exhibit no sizeable variation with temperature. In bulk crystals the LST relation has been proven experimentally, and the dramatic increase of $\epsilon(0)$ is directly related to the soft-mode behaviour. The decrease of the soft-mode eigenfrequency with temperature in bulk STO suggests a T_c of 32 K and causes $\epsilon(0)$ to increase to values above 20,000 (ref. 5), although zero-point quantum fluctuations of Ti ions prevent a ferroelectric phase transition from occurring⁷. For BST, the most commonly used ferroelectric material for DRAM applications, T_c can be adjusted by the Ba/Sr ratio to up to approximately 130 °C and $\epsilon(0)$ can be as high as 15,000 (ref. 8).

This high dielectric constant should allow the production of very compact capacitor structures. In reality, however, much lower values have been reported in thin films. For example, a dielectric constant of 150 was observed for a 24-nm-thick polycrystalline BST film in a DRAM device structure³, while a value of $\epsilon(0)$ of ~250 was found in an epitaxial 25-nm-thick STO film⁹. A widely accepted

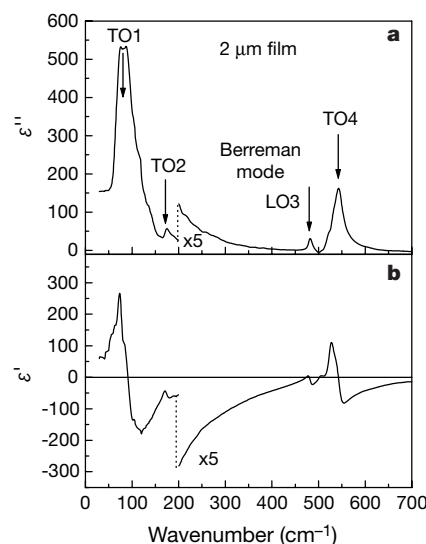


Figure 1 The effective dielectric function for a 2-μm-thick STO film at 200 K. **a**, The real part, $\epsilon'(\omega)$, and **b**, the imaginary parts, $\epsilon''(\omega)$, were measured by FT-IR ellipsometry using synchrotron radiation. From the spectra we determined the eigenfrequencies of the optical phonons by fitting to the factorized form of the dielectric function⁵. The position of the optical phonons obtained from the best fit are marked with arrows.

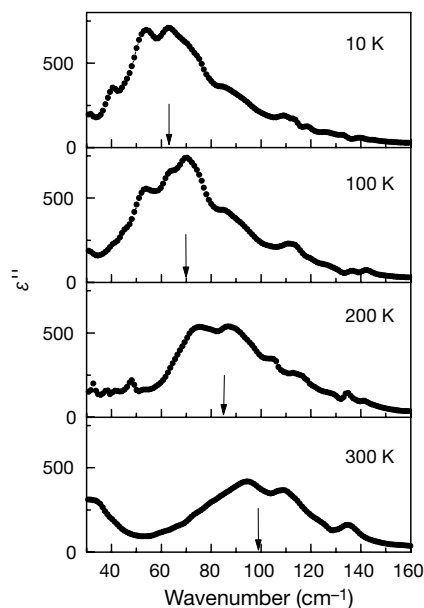


Figure 2 The imaginary part of the effective dielectric function, $\epsilon''(\omega)$, for the 2- μm -thick STO film at different temperatures. The positions of the soft-mode phonon are determined by the best fit to the factorized form of the dielectric function, and are marked with arrows.

view is that this reduction of $\epsilon(0)$ is caused by a dead layer at the ferroelectric–electrode interface which has a very low dielectric constant^{2,3,10}. Such a dead layer may arise from oxygen interdiffusion, chemical reaction, structural defects, Schottky barriers at the interfaces, or the electric-field penetration into the metal electrodes. Using a lattice dynamical approach to consider the interruption of dipole–dipole interaction by the interface, Zhou and Newns further pointed out that this dead-layer effect is intrinsic². Applying the dead-layer model to STO films of various thicknesses, we have derived a temperature-dependent dielectric constant for the volume of the film material which is only about 1,000 at its maximum; that is, still well below the value of $\epsilon(0)$ in bulk single crystals⁹. A fundamental understanding of this effect demands lattice dynamic studies, in particular on the soft-mode behaviour in thin ferroelectric films. An important question in this context is whether or not the LST relation is maintained in the thin films. Because it is rather difficult to measure the optical properties of transparent thin films, there is no report that addresses this issue. Recently, we have used a metal-oxide bilayer Raman scattering technique in STO thin films and observed strong optical-phonon peaks, indicating a lowering of crystal symmetry¹¹. However, the soft mode was not detected in the conventional Raman measurements, and an infrared spectroscopic measurement would be more suitable for detecting the infrared-active optical phonons.

We performed the far-infrared ellipsometry measurements at the National Synchrotron Light Source (NSLS) at the Brookhaven National Laboratory. This newly developed technique combines ellipsometry, which directly measures the complex dielectric function without relying on the Kramers–Kronig transformation, with Fourier-transform infrared (FTIR) spectroscopy, which allows a high throughput and multiplexing. Combined with the high brightness of the synchrotron radiation, far-infrared ellipsometry provides a powerful capability to measure vibrational properties with high reliability and accuracy. The experimental set-up and data-processing procedure have been described previously¹². The complex dielectric function was measured in the frequency range 30–700 cm^{-1} as the temperature was varied from 5 to 300 K. The samples were prepared by pulsed laser deposition with the same bilayer structure as those used in the metal-oxide bilayer Raman scattering experiments¹¹. During the deposition, the substrate was

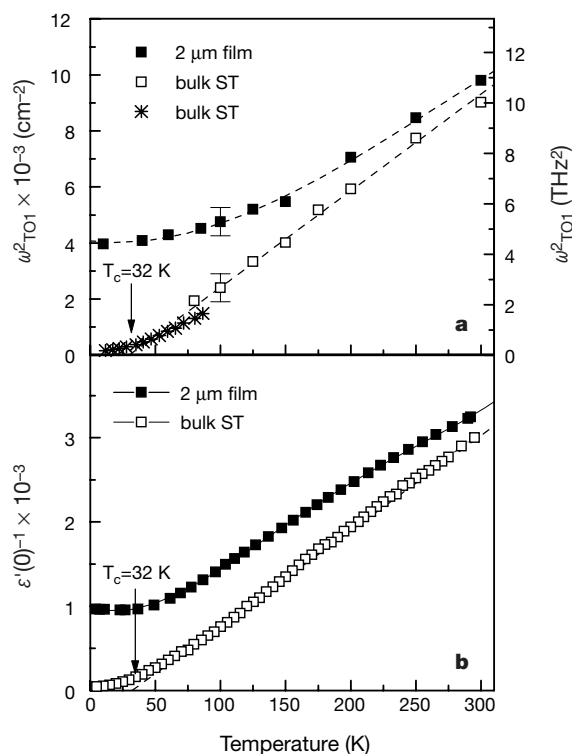


Figure 3 Comparison of the soft-mode frequencies and dielectric constants of bulk and thin-film STO. **a**, The square of the soft-mode TO1 phonon frequency versus temperature for the 2- μm -thick STO film (filled squares) and STO single crystal. The single-crystal data are from our measurement (open squares) and the hyper-Raman result of Vogt¹⁵ (stars). **b**, The inverse dielectric constant versus temperature for the 2- μm -thick STO film (filled squares) and an STO single crystal (open squares). The hardening of the soft mode in the thin film clearly correlates to a lower static dielectric constant as predicted by the LST relation.

heated to 720 °C in an oxygen pressure of 100 mtorr, and the as-deposited film was then cooled to room temperature in 400 torr of oxygen. STO films with thicknesses of 0.5, 1 and 2 μm were deposited on a 0.35- μm conducting oxide SrRuO_3 (SRO) buffer layer, which screens the optical signal from the single-crystal STO substrate. The entire bilayer structure was epitaxially grown with high crystalline quality. X-ray diffraction indicated that the STO films have a cubic structure and transform to a tetragonal structure at about 125 K, a slightly higher temperature than in the bulk crystals (~ 105 K). The static dielectric constant $\epsilon(0)$ was measured at 1 kHz using an LCR meter in a parallel-plate capacitor configuration, formed by evaporating a gold electrode onto the bilayer sample. The low-frequency loss tangent at room temperature is a few times 10^{-3} , indicating the high quality of the films.

The real and imaginary parts of the effective dielectric function, $\epsilon'(\omega)$ and $\epsilon''(\omega)$, for the 2- μm -thick STO film at 200 K are shown in Fig. 1. The soft TO1 phonon mode is clearly observed along with other TO modes, and their positions are marked in the figure. We also performed measurements on an SRO buffer layer, and ensured that the substrate signal was indeed blocked by the metallic SRO. The optical response of the SRO film compares very well to the previously reported data¹³. A detailed discussion on the correction procedure for the SRO buffer layer will be published elsewhere. The spectra from the 0.5- and 1- μm STO films showed the same general features as those in Fig. 1, except that the spectral weight of the STO–SRO interface-related Berreman mode¹⁴ decreases with increasing film thickness.

The $\epsilon''(\omega)$ spectrum between 30 and 160 cm^{-1} , which includes the peak corresponding to the soft-mode TO1 phonon, is shown as a function of temperature in Fig. 2 for the 2- μm STO film. The TO1

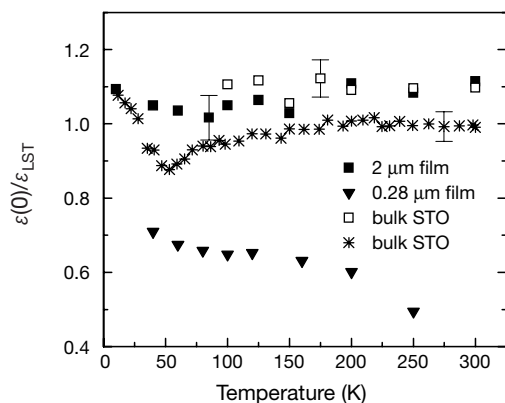


Figure 4 The agreement between the LST relation and the measured $\epsilon(0)$ in 2- μm -thick STO film. The data in Fig. 3 are replotted as the ratio $\epsilon(0)/\epsilon_{\text{LST}}$ versus temperature, where $\epsilon(0)$ is the measured value and ϵ_{LST} is the theoretical value based on the LST relation and calculated using the experimental values of the phonon frequencies. The ratio is nearly temperature independent and close to 1 for STO single crystals and the 2- μm film, indicating that the LST relation is maintained with an accuracy of about 10%. The result for a 0.28- μm -thick STO film by Fedorov *et al.*¹⁶ is also included (triangles), which shows a lower and temperature-dependent $\epsilon(0)/\epsilon_{\text{LST}}$ ratio because $\epsilon(0)$ is much smaller in the thinner film.

phonon peak is broad due to phonon damping and inhomogeneities in the films. The full-width at half-maximum (FWHM) of the soft mode in the film is about 40 cm^{-1} ; this is wider than the 25 cm^{-1} found in bulk crystals^{5,15}. The additional broadening is probably the consequence of structural defects in the STO thin film, such as oxygen vacancies¹¹. To determine the soft-mode phonon frequency, the experimental spectra are fitted to the factorized form of the dielectric function that takes into account the contribution of the phonon broadening⁵. The positions of the soft mode from the best fit are marked by the arrows in Fig. 2. As shown in this figure, the eigenfrequency of the soft mode decreases as the temperature is lowered. But in clear contrast to the bulk crystals—where the eigenfrequency of the TO1 phonon mode saturates at 13 cm^{-1} at low temperature¹⁵—in our STO thin film the TO1 eigenfrequency remains fairly high and saturates at 62 cm^{-1} . The eigenfrequencies of the other LO and TO phonon modes are rather similar in the thin film and the single crystals.

To compare our results with the theoretical prediction based on the LST relation, Fig. 3 shows the temperature dependence of the square of the soft-mode frequency, ω_{TO1}^2 , and the inverse dielectric constant, $1/\epsilon(0)$, as obtained from the low-frequency dielectric measurement for the STO film and an STO single crystal. The soft-mode frequency data for the single crystal are from our ellipsometric measurement and the hyper-Raman result by Vogt¹⁵. At room temperature, the values of ω_{TO1}^2 are about the same for the thin film and the single crystal, but when the temperature is lowered the softening of the TO1 mode is much weaker in the thin film. Because the eigenfrequencies of all the other LO and TO phonons change only slightly with temperature and are the same as those in the bulk STO, according to the LST relation $\epsilon(0)$ should depend mostly on the eigenfrequency of the soft TO1 mode. Such a trend is indeed observed: the harder soft mode in the thin film clearly correlates to a lower static dielectric constant. In Fig. 4, the ratio of the measured $\epsilon(0)$ to the theoretical value ϵ_{LST} , calculated according to equation (1) with the experimental phonon frequencies and the parameter $\epsilon(\infty) = 5.6$ (ref. 16), is plotted as a function of temperature. In the entire temperature range $\epsilon(0)/\epsilon_{\text{LST}}$ remains almost constant and close to 1 for the STO crystal and the 2- μm films. (The slight deviation from 1 at low temperatures is due to the soft-mode splitting into the A_{2u} and E_u symmetry components. Averaged soft-mode frequencies are used in Fig. 4. The low-temperature-film

values are obtained with the help of our recent electric-field-induced Raman scattering measurement²². This demonstrates that, as in the bulk crystal, the LST relation between the measured eigenfrequencies of the optical-phonon modes and the static dielectric constant is maintained with an accuracy of better than 10% in our 2- μm STO thin films.

We have previously shown that the dead layer correction to $\epsilon(0)$ is negligibly small for 2- μm -thick STO films⁹. The large STO film thickness is central to the observed agreement with the LST relation. Also shown in Fig. 4 is the $\epsilon(0)/\epsilon_{\text{LST}}$ ratio for a 0.28- μm -thick STO film obtained by Fedorov *et al.*¹⁶. Although a similar soft-mode hardening behaviour was obtained by a far-infrared transmittance experiment, the LST relation was not fulfilled because the measured static dielectric constant was much smaller than in our films. This reconfirms that the $\epsilon(0)$ reduction due to the dead-layer effect is more important for smaller film thickness.

From the LST relation, we obtain

$$\frac{\omega_{\text{TO1, film}}^2}{\omega_{\text{TO1, bulk}}^2} = \frac{\epsilon_{\text{bulk}}(0)}{\epsilon_{\text{film}}(0)}$$

where $\omega_{\text{TO1, film}}$ and $\omega_{\text{TO1, bulk}}$ denote the eigenfrequencies of the soft TO1 mode in the thin film and the bulk crystal, respectively. Our result confirms this relation, and suggests that besides the dead-layer effect, the lower dielectric constant in STO thin films is due to their fundamental lattice dynamical properties—specifically, the soft-mode hardening. The physical mechanisms for the soft-mode hardening now become the central issue. It is rather surprising that our highly crystalline thick film exhibits soft-mode behaviour similar to that shown by the very thin, polycrystalline film investigated by Fedorov *et al.*¹⁶, which seems to suggest a mechanism related to the reduced dimensionality. On the other hand, it has been shown that much higher dielectric constants can be obtained in STO films^{17,18} and that the epitaxial growth mode has a marked influence on their dielectric properties¹⁸, which suggests a critical role of strain and defects. An analysis of our recent electric-field-induced Raman scattering results in STO thin films indicates that the local polar regions surrounding the oxygen vacancies contribute to the soft-mode hardening.

Our result has implications for the theories of dielectric response in ferroelectric thin films. The phenomenological theory that attributes the reduction of permittivity to the interface-modified depolarization field¹⁹ should be reconciled with the soft-mode behaviour in thin films. It also shows that the microscopic theory of the intrinsic dead-layer effect² needs to be modified to extend it beyond the soft-mode hardening in the interfacial layer. The result we report here could also be relevant to investigations of other ferroelectric materials, and to the ‘size effect’ on ferroelectric properties in polycrystalline ceramics. For example, in polycrystalline BST thin films, the reduction of dielectric constant has been attributed to interfacial dead-layer effect, cation nonstoichiometry, and in-plane stress²⁰. In fine-grained BaTiO_3 , the small grain size and stress are closely linked, and cause an increase in the dielectric constant²¹. Such size effects will no doubt influence soft-mode behaviour in thin films and fine particles, with significant implications for high- $\epsilon(0)$ applications of ferroelectric thin films. □

Received 31 August; accepted 20 December 1999.

- Kotecki, D. E. *et al.* (Ba,Sr)TiO₃ dielectrics for future stacked-capacitor (DRAM). *IBM J. Res. Dev.* **43**, 367–382 (1999).
- Zhou, C. & News, D. M. Intrinsic dead layer effect and the performance of ferroelectric thin film capacitors. *J. Appl. Phys.* **82**, 3081–3088 (1997).
- Basceri, C. *et al.* The dielectric response as a function of temperature and film thickness of fiber-textured (Ba, Sr)TiO₃ thin films grown by chemical vapor deposition. *J. Appl. Phys.* **82**, 2497–2504 (1997).
- Shirane, G. & Yamada, Y. Lattice-dynamical study of the 110 °K phase transition in SrTiO₃. *Phys. Rev.* **177**, 858–863 (1969).
- Servoin, J. L. *et al.* Infrared dispersion in SrTiO₃ at high temperature. *Phys. Rev. B* **22**, 5501–5506 (1980).

6. Cochran, W. Crystal stability and the theory of ferroelectricity. *Adv. Phys.* **9**, 387 (1960).
7. Müller, K. A. & Burkard, H. SrTiO₃: an intrinsic quantum paraelectric below 4 K. *Phys. Rev. B* **19**, 3593–3602 (1979).
8. Mitsui, T. & Nomura, S. in *LANDOLT-BÖRNSTEIN Numerical Data and Functional Relationships in Science and Technology* (eds Hellwege, K. H. & Hellwege, A. M.) Vol. 16(a) 328–359 (Springer, Berlin, 1981).
9. Li, H. C. *et al.* Thickness dependence of dielectric loss in SrTiO₃ thin films. *Appl. Phys. Lett.* **73**, 464–466 (1998).
10. Black, C. T. & Welser, J. J. Electric-field penetration into metals: Consequences for high-dielectric-constant capacitors. *IEEE Trans. Electron Devices* **46**, 776–780 (1999).
11. Sirenko, A. A. *et al.* Observation of the first-order Raman scattering in SrTiO₃ thin films. *Phys. Rev. Lett.* **82**, 4500–4503 (1999).
12. Henn, R. *et al.* Far infrared ellipsometry using synchrotron radiation: the out-of-plane response of La₂Sr₂CuO₄ thin solid films. **313**, 642–648 (1998).
13. Kostic, P. *et al.* Non-Fermi-liquid behavior of SrRuO₃: Evidence from infrared conductivity. *Phys. Rev. Lett.* **81**, 2498–2501 (1998).
14. Berreman, D. W. Infrared absorption at longitudinal optic frequency in cubic crystal films. *Phys. Rev.* **130**, 2193–2198 (1963).
15. Vogt, H. Refined treatment of the model of linearly coupled anharmonic oscillators and its application to the temperature dependence of the zone-center soft-mode frequencies of KTaO₃ and SrTiO₃. *Phys. Rev. B* **51**, 8046–8059 (1995).
16. Fedorov, I. *et al.* Far-infrared spectroscopy of a SrTiO₃ thin film. *Ferroelectrics* **208**, 413–427 (1998).
17. Fuchs, D. *et al.* High dielectric constant and tunability of epitaxial SrTiO₃ thin film capacitors. *J. Appl. Phys.* **85**, 7362–7369 (1999).
18. Lippmaa, M. *et al.* Step-flow growth of SrTiO₃ thin films with a dielectric constant exceeding 10⁴. *Appl. Phys. Lett.* **74**, 3543–3545 (1999).
19. Vendik, O. G. & Ter-Martirosyan, L. T. Size effect in layered structures, ferroelectric/normal metal and ferroelectric/high T_c superconductor. *Sov. Phys. Solid State* **36**, 1778–1781 (1994).
20. Streiffer, S. K. *et al.* Ferroelectricity in thin films: The dielectric response of fiber-textured (Ba_{0.8}Sr_{0.2})Ti_{1-x}O_{3-x/2} thin films grown by chemical vapor deposition. *J. Appl. Phys.* **86**, 4565–4575 (1999).
21. Frey, M. H. & Payne, D. A. Grain-size effect on structure and phase transformations for barium titanate. *Phys. Rev. B* **54**, 3158–3168 (1996).
22. Akimov, I. A. *et al.* Electric field-induced soft-mode hardening in SrTiO₃ films. *Phys. Rev. Lett.* (submitted).

Acknowledgements

We thank M. Cardona, H. Vogt, L. E. Cross, D. M. Newns, T. M. Shaw, S. B. Desu, J. Humlíček and J. Petzelt for discussions; G. P. Williams and L. Carr for technical support at the U4IR and U12IR beamlines at NSLS at Brookhaven National Laboratory, a US DOE supported user facility; and I. A. Akimov for the Raman scattering measurements. This work was partially supported by the NSF, DOE and DARPA.

Correspondence and requests for materials should be addressed to A.A.S. (e-mail: sirenko@phys.psu.edu).

Extended surface chirality from supramolecular assemblies of adsorbed chiral molecules

M. Ortega Lorenzo*, C. J. Baddeley*, C. Muryn*† & R. Raval*

* Leverhulme Centre for Innovative Catalysis and Surface Science Centre, Department of Chemistry, University of Liverpool, Liverpool L69 7ZD, UK

The increasing demand of the chemical and pharmaceutical industries for enantiomerically pure compounds has spurred the development of a range of so-called 'chiral technologies' (ref. 1), which aim to exert the ultimate control over a chemical reaction by directing its enantioselectivity². Heterogeneous enantioselective catalysis^{3–5} is particularly attractive because it allows the production and ready separation of large quantities of chiral product while using only small quantities of catalyst. Heterogeneous enantioselectivity is usually induced by adsorbing chiral molecules onto catalytically active surfaces^{1,3–7}. A mimic of one such catalyst is formed by adsorbing (*R,R*)-tartaric acid molecules on Cu(110) surfaces: this generates a variety of surface phases, of which only one is potentially catalytically active⁸, and leaves the

question of how adsorbed chiral molecules give rise to enantioselectivity. Here we show that the active phase consists of extended supramolecular assemblies of adsorbed (*R,R*)-tartaric acid, which destroy existing symmetry elements of the underlying metal and directly bestow chirality to the modified surface. The adsorbed assemblies create chiral 'channels' exposing bare metal atoms, and it is these chiral spaces that we believe to be responsible for imparting enantioselectivity, by forcing the orientation of reactant molecules docking onto catalytically active metal sites. Our findings demonstrate that it is possible to sustain a single chiral domain across an extended surface—provided that reflection domains of opposite handedness are removed by a rigid and chiral local adsorption geometry, and that inequivalent rotation domains are removed by successful matching of the rotational symmetry of the adsorbed molecule with that of the underlying metal surface.

The hydrogenation of β -ketoesters over Ni catalysts that have been modified with (*R,R*)-tartaric acid ((*R,R*)-TA) is one the most successful heterogeneous enantioselective reactions, producing enantiomeric excesses of >90% of the hydrogenated *R*-product^{4,6,7}. The enantioselectivity can be switched to yield the *S*-enantiomer product by simply using the mirror-image (*S,S*)-tartaric acid as a modifier⁶. When a single molecular layer of (*R,R*)-tartaric acid is adsorbed on Cu(110), only the ordered phase created at low coverages and temperatures in excess of 350 K is capable of subsequent adsorption of the reactant molecule, methylacetoacetate, which represents the simplest β -ketoester. Surface infrared data⁸ indicate that this active phase consists solely of the bitartrate form in which both acid groups have deprotonated to yield carboxylate moieties. The bitartrate molecule straddles the Cu rows so that both carboxylate ends are positioned above the preferred⁹ two-fold short-bridge adsorption site, and the oxygen atoms are equidistant from the surface⁸ (Fig. 1A). This adsorption mode leaves the C2–C3 bond almost parallel to the surface, with the C–OH groups extending out from the molecular skeleton, projecting a large component parallel to the surface.

Here we present a detailed study of the two-dimensional ordering of this active phase, which gives rise to a low-energy electron diffraction (LEED) pattern (Fig. 1B) that can be indexed in matrix notation as (9 0, 1 2) (see Fig. 1B legend for details of this nomenclature). Scanning transmission microscopy (STM) images of this phase (Fig. 1C) confirm the presence of long-range order, yielding an oblique (9 0, 1 2) repeat surface unit mesh, with the dimensions (Å) 23.04 × 7.68, $\alpha = 19.47^\circ$. Each unit cell contains three molecules, resulting in a fractional coverage of 1/6 with respect to the number density of surface metal atoms. The STM images reveal that the bitartrate molecules are self-assembled in rows of three, each row stacking in parallel with others to form long chains. These long chains are aligned along the $\langle 1\bar{1}4 \rangle$ surface direction, which does not coincide with any of the symmetry directions present on the Cu(110) surface. As a result, this molecular growth direction has the effect of destroying all the symmetry elements of the underlying metal, leading directly to the creation of a chiral surface that is non-superimposable on its mirror image. This observation suggests that cooperative¹⁰ behaviour of the adsorbed bitartrate molecules plays a central role in bestowing chirality, or 'handedness', to the achiral metal surface. We attribute this self-assembly to the close proximity of the α -hydroxy groups on neighbouring bitartrate molecules, leading to intermolecular hydrogen-bonding interactions that extend across the surface. The growth direction of this supramolecular assembly is dictated solely by the conformation of the α -hydroxy groups on the adsorbed bitartrate species (Fig. 1A); in this case, the optimum interaction leads to a $\langle 1\bar{1}4 \rangle$ chain alignment. We note that in order to restrict supramolecular growth along one particular direction, the locations of the hydroxy groups need to be constrained in space. The bitartrate form meets this requirement by its two-point attachment

† Permanent address: IRC for Surface Science, University of Manchester, Manchester M13 9PL, UK.

to the surface via its carboxylate groups, which imparts a rigid and defined adsorption geometry.

We also investigated the behaviour of the mirror-image modifier, (*S,S*)-tartaric acid, in order to verify that directed supramolecular assembly is the pivotal factor in enabling modifiers to bestow chirality. The active (*S,S*)-TA bitartrate phase was created under identical conditions used for (*R,R*)-TA. LEED patterns for this (*S,S*)-TA phase confirm that a chiral surface has been created. Significantly, however, the chirality of the surface has now been switched, as illustrated by the STM images (Fig. 2): the (9 0, -1 2) phase of (*S,S*)-TA on Cu(110) is a true mirror image of the (9 0, 1 2) phase obtained for (*R,R*)-TA. That is, the (*S,S*)-TA adsorbed layer shows chain growth in the $\langle 1\bar{1}4 \rangle$ direction, and the alignment of individual molecules in each row of three has also been switched to the mirror orientation.

This chiral switching can be explained in terms of molecular structure and its effect on supramolecular assembly. Both TA enantiomers possess defined and rigid adsorption geometries arising from their two-point bonding with the metal surface. The only difference between them is the spatial orientation of their OH groups (Fig. 3a). For (*R,R*)-TA, the positions of the OH groups and the directionality of intermolecular hydrogen-bonds leads to a molecular chain alignment along the $\langle 1\bar{1}4 \rangle$ direction. For (*S,S*)-TA, the mirror positioning of the hydroxy groups now naturally forces chain growth along the mirror $\langle 114 \rangle$ direction (Fig. 3a), explaining why the mirror unit mesh is formed. The effect of changing the enantiomer form is also manifested in the mirror alignment of molecules in each row of three. This suggests that intermolecular hydrogen-bonding interactions are not just confined along the long-chain direction but, rather, an extended hydrogen-bonding network must exist that weaves the three chains together. In Fig. 3b we propose such a network where, again, the mirror positions of the OH groups in the two enantiomers would lead to a switch in supramolecular assembly directions. This work clearly demon-

strates that the role of supramolecular assemblies in inducing chirality^{11–13} can be extended to the adsorbed monolayer. Such hydrogen-bonding networks also contribute towards making the chiral surfaces created here very robust and stable up to temperatures of >200 °C.

The models proposed here have a number of implications for enantioselective heterogeneous catalysis^{1,3–7}. First, when simple modifiers are used, it seems that the active site is defined not by single adsorbed chiral molecules, but by supramolecular assemblies.

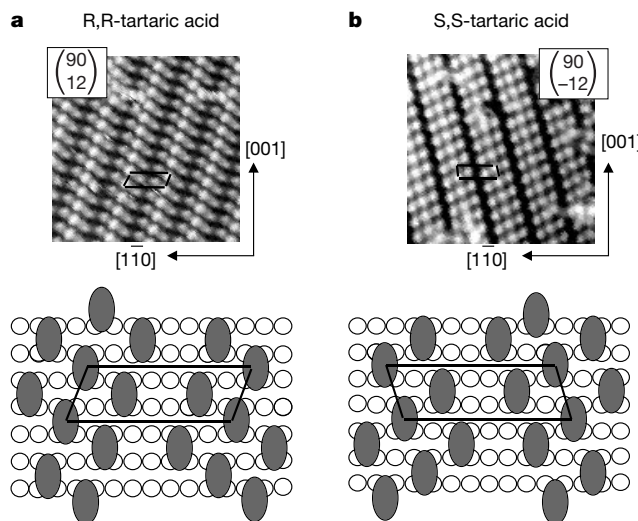


Figure 2 STM images showing mirror chiral surfaces created by (*R,R*)- and (*S,S*)-TA adsorbed on Cu(110). Shown are 108 Å × 108 Å STM images of the (9 0, 1 2) (*R,R*)-TA bitartrate phase (**a**; $V_{\text{tip}} = -1.7$ V; $I_{\text{tip}} = 1.18$ nA) and the (9 0, -1 2) (*S,S*)-TA bitartrate phase (**b**; $V_{\text{tip}} = -2.73$ V; $I_{\text{tip}} = 1.02$ nA). Below each STM image is a schematic diagram displaying the position of the molecular features observed relative to the Cu(110) surface.

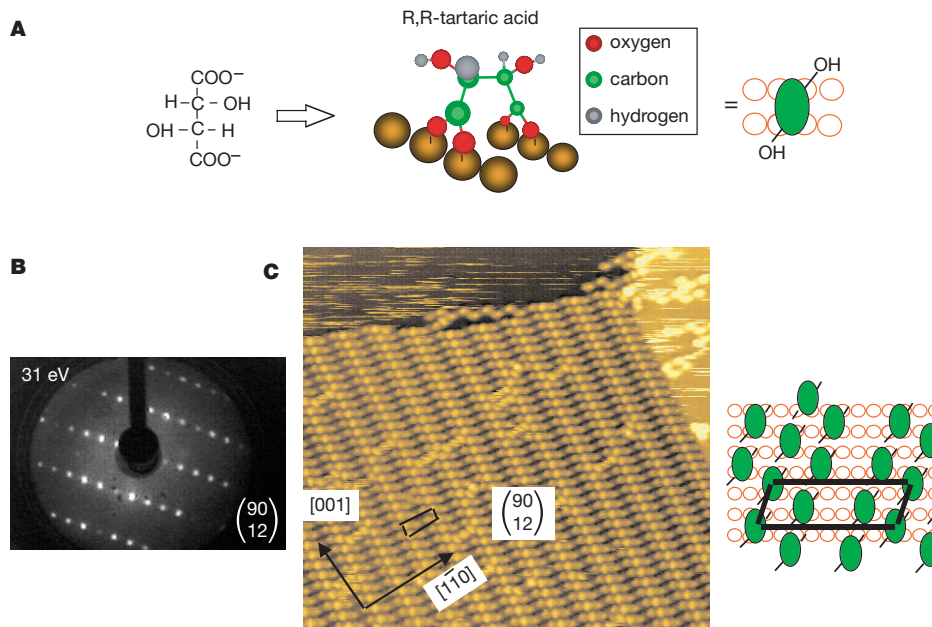


Figure 1 The active chiral phase created by (*R,R*)-TA on Cu(110). Shown is local adsorbate geometry and extended two-dimensional ordering. **A**, Schematic diagram showing the local bonding and orientation of the adsorbed bitartrate species on a Cu(110) surface. **B**, Low-energy electron diffraction (LEED) pattern observed at 31 eV for the ordered (9 0, 1 2) bitartrate phase of (*R,R*)-TA on Cu(110). **C**, Scanning tunnelling microscope (STM) image (150 Å × 200 Å) showing the ordered (9 0, 1 2) bitartrate phase formed by (*R,R*)-tartaric acid adsorption on Cu(110) at 405 K ($V_{\text{tip}} = -1.7$ V; $I_{\text{tip}} = 1.18$ nA).

The two-dimensionally ordered structures are described in matrix notation with reference to **a** and **b**, the unit vectors along $\langle 1\bar{1}0 \rangle$ and along the $\langle 001 \rangle$ directions on the Cu(110) surface. For convenience, the matrix is represented throughout in the text as (G_{11} G_{12} , G_{21} G_{22}). STM and LEED experiments were performed in ultrahigh vacuum chambers. Atomically clean Cu(110) surfaces were exposed to tartaric acid by direct sublimation of the 99% pure compounds into the chamber.

For the tartaric acid phases studied here, we find that between the rows of adsorbed molecules lie empty chiral 'channels' that expose bare metal atoms, which are essential to bind the incoming reactant molecules. We propose these nanosized chiral spaces may constitute the actual active enantioselective site in which the reactant molecule docks in a preferential orientation, forcing the hydrogenation to occur at one reactant face only. Switching the chirality of the modifier switches the handedness of both the adsorbed surface phase and the channels, forcing docking to occur in the opposite sense, subjecting the other reactant face to hydrogenation, thus switching enantioselectivity from the *R*- to the *S*-reaction product⁶. Studies of the real catalyst systems show that enantioselectivity actually increases with metal catalyst particle size⁶, the optimal value being 100–200 Å. This result suggests that enantioselective catalysis occurs on the extended metal terraces, and that the underlying reason for increased enantiomeric yields is that facet sizes above a critical value are required to enable the creation of supramolecular modifier assemblies with chiral channels.

So far, only restricted examples of local chirality at surfaces have been demonstrated^{14,15}, but the STM and LEED images of the (9 0, 1 2) and (9 0, -1 2) TA phases in the present work show that it is possible to create chiral surfaces sustaining a single handedness across the entire surface. However, the adsorption of chiral molecules at surfaces does not automatically lead to the creation of chiral phases, as demonstrated by the various achiral phases created by (*R,R*)-TA on Cu(110) (ref. 8). We now therefore consider some general factors that constrain the expression of chirality at an extended two-dimensional surface.

Of the 230 space groups that exist in three-dimensional systems, only 17 are allowed in two-dimensional systems^{16,17}. Further elimination of mirror and glide symmetry elements leaves only 5 space groups that are consistent with two-dimensional chiral systems, described in Schoenflies notation¹⁶ as: C_1 , C_2 , C_3 , C_4 and C_6 . Additionally, to ensure chiral integrity across the entire surface, all

chiral domains formed must be equivalent and exhibit the same handedness. However, during adsorption, molecules arrive randomly and occupy energetically equivalent adsorption geometries with equal probability, forming nucleation meshes that define local domain structure. Generally, translational, reflection and rotational surface domains¹⁸ can be created. Whereas translational domains can be tolerated, as translational symmetry operations do not destroy the 'handedness' of the adsorbed layer, reflection and rotational domains could potentially allow different two-dimensional surface stereoisomers to co-exist. This problem could be eliminated in two ways: either ensuring that all domains formed are equivalent, or ensuring that inequivalent domains are physically disallowed. For example, inequivalent rotational domains occur when the overlayer surface mesh possesses a lower symmetry than that of the clean substrate. So, to ensure that all rotational domains are equivalent, the rotational symmetry of the adsorbed layer unit mesh must match that of the bare surface, so all equally probable adsorption geometries result in an identical nucleation point. This would mean that on f.c.c.(110), f.c.c.(100) and f.c.c.(111) metal surfaces, the manifestation of a single chiral domain is restricted to overlayer unit meshes with space groups C_2 , C_4 and C_3/C_6 , respectively.

Extended chirality is particularly threatened by reflection domains that would lead to the coexistence of mirror-image domains. The chirality of the surface unit mesh means that reflection domains are necessarily inequivalent, so this problem can only be eliminated if the creation of the reflection domain is physically impossible. This criterion is met if the adsorbed entity is itself chiral, ensuring that its mirror image cannot be physically realized at the surface.

The active bitartrate phase discussed here fulfills all these requirements. The chirality of the molecule and its two-point rigid bonding at the surface leads to a C_2 point group, eliminating the possibility of reflection domains. In addition, directed supramolecular assembly of the adsorbed bitartrate units produces a unique growth direction, creating an oblique chiral unit mesh belonging to the C_2 space group. This unit mesh matches the rotational symmetry of the Cu(110) surface, ensuring that all rotational domains are equivalent, so that a single chiral domain is sustained across the entire surface. □

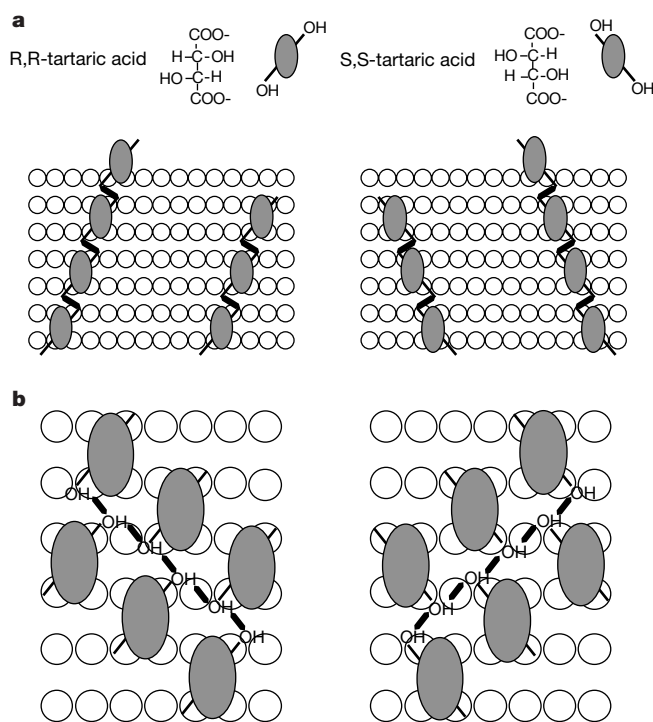


Figure 3 The spatial alignment and intermolecular hydrogen-bonding interactions of the α -hydroxy groups on the two enantiomers of tartaric acid. Left column, (*R,R*)-TA; right column, (*S,S*)-TA. The short, thick lines show extended hydrogen-bonding interactions: in **a**, these interactions dictate the direction of long chain growth; in **b**, they 'weave together' the three molecular chains.

Received 2 August 1999; accepted 20 January 2000.

- Sheldon, R. A. *Chirotechnology* (Marcel Dekker, New York, 1993).
- Brown, J. M. & Davies, S. G. Chemical asymmetric synthesis. *Nature* **342**, 631–636 (1989).
- Baiker, A. & Blaser, H. U. in *Handbook of Heterogeneous Catalysis* Vol. 5 (eds Ertl, G. H., Knoezinger, H. & Weinheim, J.) 2422–2430 (VCH, New York, 1997).
- Blaser, H. U. Enantioselective synthesis using chiral heterogeneous catalysts. *Tetrahedron: Asymmetry* **2**, 843–866 (1991).
- Wan, T. A. & Davies, M. E. Design and synthesis of a heterogeneous asymmetric catalyst. *Nature* **370**, 449–450 (1994).
- Tai, A. & Harada, T. in *Tailored Metal Catalysts* (ed. Iwasawa, Y.) 265–324 (Reidel, Dordrecht, 1986).
- Izumi, Y. Modified Raney nickel catalyst: Heterogeneous enantio-differentiating catalyst. *Adv. Catal.* **32**, 215–270 (1983).
- Ortega Lorenzo, M., Hag, S., Murray, P., Raval, R. & Baddeley, C. J. Creating chiral surfaces for heterogeneous catalysis: *R,R*-tartaric acid on Cu(110). *J. Phys. Chem.* **103**, 10661–10669 (1999).
- Woodruff, D. P. *et al.* The structure of the formate species on copper surfaces: new photoelectron diffraction and SEXAF data reassessed. *Surf. Sci.* **201**, 228–244 (1988).
- Lehn, J.-M. *Supramolecular Chemistry, Concepts and Perspectives* (VCH, New York, 1995).
- Riviera, J. M., Martin, T. & Rebek, J. Jr Chiral spaces: dissymmetric capsules through self-assembly. *Science* **279**, 1021–1023 (1998).
- Prins, L. J. *et al.* Complete asymmetric induction of supramolecular chirality in a hydrogen-bonded assembly. *Nature* **398**, 498–501 (1999).
- MacGillivray, L. R. & Atwood, J. L. A chiral spherical molecular assembly held together by 60 hydrogen bonds. *Nature* **389**, 469–472 (1997).
- Lopinski, G. P., Moffatt, D. J., Wayner, D. D. M. & Wolkow, R. A. Determination of the absolute chirality of individual adsorbed molecules using the scanning tunnelling microscope. *Nature* **392**, 909–911 (1998).
- Viswanathan, R., Zasadzinski, J. A. & Schwartz, D. K. Spontaneous chiral symmetry breaking by achiral molecules in a Langmuir-Blodgett film. *Nature* **368**, 440–443 (1994).
- International Tables for X-Ray Crystallography* Vol. 1 (eds Henry, N. F. M. & Lonsdale, K.) (Kynoch, Birmingham, 1965).
- Richardson, N. V. & Sheppard, N. in *Vibrational Spectroscopy of Molecules on Surfaces* (eds Yates, J. T. Jr & Madey, T. E.) 1–48 (Plenum, New York, 1987).
- Woodruff, D. P. & Delchar, T. A. *Modern Techniques of Surface Science* (Cambridge University Press, Cambridge, 1986).

Acknowledgements

We thank the Leverhulme Centre for Innovative Catalysis for a studentship (M.O.L.), and also thank the EPSRC for equipment grants and a post-doctoral fellowship (C.M.).

Correspondence and requests for materials should be addressed to R.R. (e-mail: R.Raval@liv.ac.uk).

Bonding and reactivity at oxide mineral surfaces from model aqueous complexes

Brian L. Phillips*, William H. Casey†‡ & Magnus Karlsson†§

* Department of Chemical Engineering and Materials Science, †Department of Land, Air and Water Resources, ‡Department of Geology, University of California, Davis, California 95616, USA

The kinetic stability of oxide surfaces affects a broad range of physical phenomena, including mineral dissolution^{1–3} and sorption reactions⁴, stable-isotope fractionation⁵, and catalyst support degradation⁶. Our knowledge of the rates of these processes derives mostly from the rates of net mass transfer between the bulk solid and fluid phases. But from such data it is difficult to determine rates of elementary steps that are needed to test theoretical models. Here we determine the rates of oxygen exchange between an aqueous fluid and specific sites on the 'Al₁₃' polyoxocation—AlO₄Al₁₂(OH)₂₄(H₂O)₁₂⁷⁺—the structure of which closely resembles the surfaces of some Al-(hydr)oxide minerals in soils and catalyst supports. Extrapolation of these data to 298 K (and near pH 5.3) yields half-lives for oxygen on the complex that range from ~0.6 milliseconds for bound water to 41 seconds and 13 hours for the two distinct, but structurally similar, bridging hydroxyls. This surprisingly large range of labilities (~10⁷) indicates that reactivity is very sensitive to molecular structure. Moreover, these results indicate that well chosen aqueous complexes provide important information to relate bonding to reactivity at mineral surfaces.

The rate of oxygen exchange between oxide minerals and aqueous fluids is the fundamental information about bond-breaking needed to understand more complicated, multi-step processes, such as mineral dissolution and sorption. However, no method exists for measuring *in situ* the rates of elementary reactions in heterogeneous systems. Some guidance comes from studies of large dissolved molecules that are amenable to kinetic measurements. Although a large body of oxygen exchange-rate data exists for inert aqueous complexes, such as those of Cr³⁺ and Mo⁴⁺ (see, for example, refs 7, 8), few polynuclear species in aqueous solution are known for the more geochemically important metals, such as Al. However, the most familiar aqueous multimer of Al, the 'Al₁₃' polyoxocation (AlO₄Al₁₂(OH)₂₄(H₂O)₁₂⁷⁺), is large and bears a striking resemblance to the structure of Al-(hydr)oxides (Fig. 1). The 'Al₁₃' molecule, an example of the ϵ -Keggin structure, comprises a central Al(O)₄ tetrahedron surrounded by 12 octahedrally coordinated Al that share edges (ref. 9). Each Al octahedron exposes one water molecule to the solution that can deprotonate at pH > 6 (ref. 10). Four different oxygens can be distinguished: an internal μ_4 -oxo, the terminal η -OH₂, and two types of μ_2 -OH. The bridging hydroxyls differ in their positions relative to the μ_4 -oxo: μ_2 -OH has one μ_4 -oxo in the *cis* position, whereas μ_2 -OH' lies adjacent to two μ_4 -oxo groups (Fig. 1a). Similarly, most Al-(hydr)oxide phases,

such as bayerite (Fig. 1b), contain edge-linked Al octahedra with surfaces that also contain terminal water (and/or hydroxyls) and paired bridging hydroxyls which interact with the solvent¹¹. 'Al₁₃' can also be viewed as a subunit of γ -alumina (Fig. 1c), a commonly used catalyst support that also contains tetrahedrally coordinated Al bonded to μ_4 -oxo groups. In addition to structural similarities, the surface charge density of 'Al₁₃', 0.32 C m⁻² (+7 charge, *r* = 0.53 nm; ref. 9) is comparable to those for minerals such as bayerite at pH values less than the point of zero net proton charge, 0.16–0.48 C m⁻² (refs 3, 12).

Rates of oxygen exchange between less labile sites on the 'Al₁₃' complex and solvent water were measured by the change in ¹⁷O NMR intensity after addition of ¹⁷O-enriched water to an isotopically normal solution. Absolute intensity changes were obtained by normalizing the integrated intensity to that from an external standard, an aqueous 0.3 M Tb³⁺ solution that gives a narrow ¹⁷O NMR peak near -100 p.p.m. (Fig. 2). We prepared monospecific solutions of 'Al₁₃' by dissolving crystals of its selenate salt [NaAlO₄Al₁₂(OH)₂₄(H₂O)₁₂(SeO₄)₂] in a BaCl₂ solution and filtering¹⁰. This solution was then mixed with an equal volume of water enriched to 30% in ¹⁷O (Isotec Inc.) and containing 0.5 M Mn(II). The presence of Mn(II) broadens the NMR peak for solvent water beyond detection¹³. These solutions exhibit a nearly constant apparent ionic strength of ~1.3 M and a narrow range of initial pH values, from 5.2 to 5.4. The pH decreased slightly (< 0.24 units)

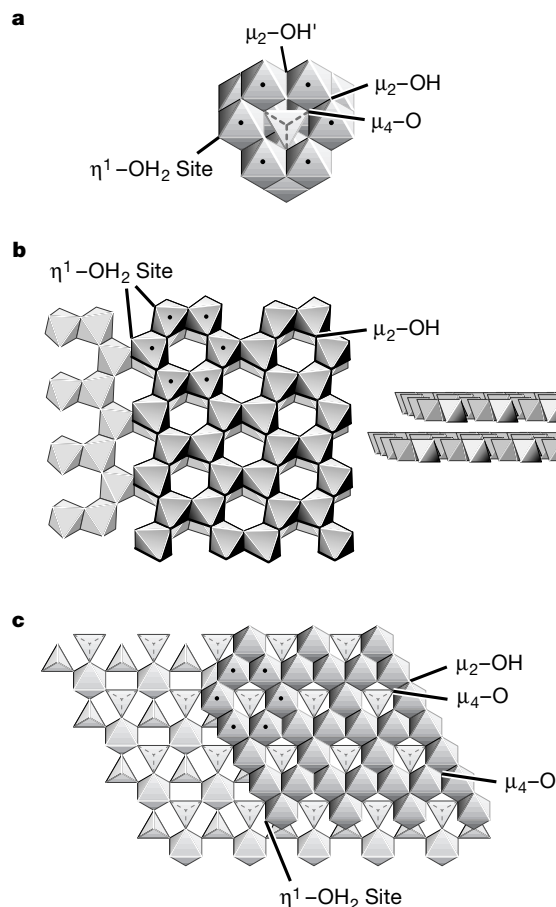


Figure 1 Polyhedral representations of structures discussed in the text. **a**, The 'Al₁₃' ϵ -Keggin-type polyoxocation (AlO₄Al₁₂(OH)₂₄(H₂O)₁₂⁷⁺); **b**, a monomolecular step on bayerite (γ -Al(OH)₃), with inset showing its layered structure; and **c**, a monomolecular step on γ -Al₂O₃. Octahedra and tetrahedra represent Al(O)₆ and Al(O)₄ bonding units, respectively. A fraction of the tetrahedral Al sites are vacant in γ -Al₂O₃. Specific types of oxygens exposed near the surface are labelled, and can undergo Brønsted acid–base reactions with the solvent.

§ Present address: Department of Inorganic Chemistry, Umeå University, S 90187 Umeå, Sweden.

during the experiment (120 to 3,000 h). Nearly all the aluminium in these solutions occurs as Al_{13} , as indicated by comparing the absolute ^{27}Al NMR intensity of the central tetrahedral Al with that of the $Al(H_2O)_6^{3+}(aq.)$ produced upon complete decomposition of the complex by acidification¹⁰.

At temperatures between 282 and 292 K, ^{17}O NMR spectra taken immediately upon addition of the $H_2^{17}O$ contain a peak at +20 p.p.m., in addition to the peak near -100 p.p.m. from the intensity standard (Fig. 2a). The total spectral intensity doubles within a few hours, with gradual development of intensity at higher chemical shifts (Fig. 2b). The rate of intensity increase then slows considerably. The total intensity asymptotically approaches a value that is approximately three times the initial intensity over times of from one day to several weeks, depending on the temperature. For each sample, spectra taken at long equilibration times (Fig. 2c) show a broad, partially resolved peak near +30 p.p.m. These spectra are well-represented by a sum of three lorentzian curves (including the intensity standard), and least-squares fits give relative intensities of 2:1 for the peaks at +30 and +20 p.p.m., respectively. Unconstrained fits to spectra taken at shorter times differ only in the intensity of the +30 p.p.m. peak and yield values for the position and width of the peak at +20 p.p.m. that do not vary significantly, although its absolute intensity decreases slightly over the course of the experiment (<8%). These observations indicate that the increase of ^{17}O NMR intensity with time can be associated with the peak near +30 p.p.m.

These experimental observations are consistent with isotopic equilibration of the hydroxyl sites by oxygen exchange with solvent:

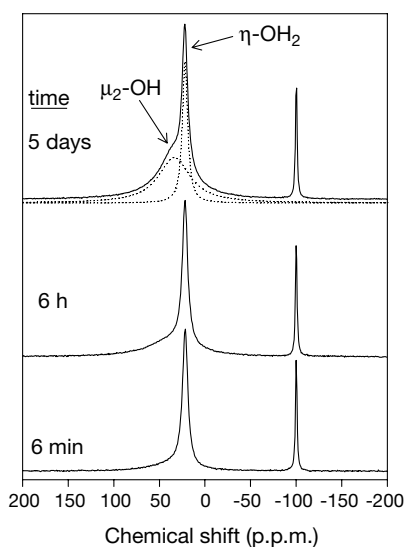
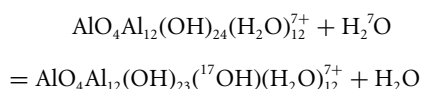


Figure 2 ^{17}O NMR spectra at 289 K of a ~ 10 mM aqueous solution of $AlO_4Al_{12}(OH)_{24}(H_2O)_{12}^{7+}$ with 0.25 M $Mn(II)$. Spectra are shown as a function of time after addition of water enriched to 30% ^{17}O . The peak near +20 p.p.m. exhibits nearly constant intensity, and is assigned to the η -OH₂ sites on the complex. Intensity increase with time is due to a peak near +30 p.p.m. that is assigned to μ_2 -OH and μ_2 -OH'. Least-squares fits of spectra obtained after long reaction times give intensity ratios $\sim 2:1$ for the peaks at 30 and 20 p.p.m. (in the top panel, components are shown as dotted lines with baseline displaced for clarity). The internal μ_4 -oxo sites are inert, and cannot be detected at natural ^{17}O abundance. Spectra are scaled to the peak at -100 p.p.m., from water in the $Tb^{3+}(aq.)$ intensity standard. Spectra represent 20,000 acquisitions by 20- μs pulses ($\sim \pi/3$) and 5-ms relaxation delays, obtained at 67.8 MHz with a Bruker Avance NMR spectrometer. Frequency scale is referenced externally to tap water.

We assign the peak at +20 p.p.m. to the η -OH₂ groups of the complex, based on its chemical shift and rapid exchange with solvent water, both of which are consistent with previous data for aqueous monomeric complexes of $Al(III)$ (refs 14, 15). The very slow decrease of intensity of this peak, and slight decrease of solution pH, are consistent with previous measurements for the decomposition rate of Al_{13} at pH 5 (half-life, $t_{1/2} > 500$ h at 298 K; refs 11, 16). Monomeric $Al(III)$ complexes do not give an observable NMR signal under these conditions¹⁷ and would not significantly affect the ^{17}O NMR results if small amounts formed as decomposition products. The near-integral increases in ^{17}O NMR intensity with time, and the fitted relative peak ratios at long times, are consistent with the stoichiometry of the complex and assignment of the peak at +30 p.p.m. to the unresolved μ_2 -OH and μ_2 -OH' sites. We observed no signal from the μ_4 -oxo. A signal for this site (+55 p.p.m.) has been observed for Al_{13} prepared directly from ^{17}O -enriched reagents¹⁸.

The change of ^{17}O NMR intensity with time (Fig. 3), associated with exchange at hydroxyl sites, $\Delta I(t) = I(t) - I(0)$ follows a simple bi-exponential kinetic law of the form:

$$\Delta I(t) = 2 - \exp[-k_{ex,a}t] - \exp[-k_{ex,b}t] \quad (2)$$

where t is the time elapsed since addition of the $H_2^{17}O$, $I(0)$ is the initial intensity of the peak at 20 p.p.m. (typically a few minutes after mixing), and $k_{ex,a} \gg k_{ex,b}$. All intensities were normalized to the standard, and those at long times were adjusted slightly for decreases in complex concentration indicated by the small decrease in the peak at 20 p.p.m. Equation (2) is a simplified form of the McKay equation for isotopic equilibration of distinct species where the initial ^{17}O content of the complex is negligible and the complex contains a small fraction of total oxygen in solution¹⁹. Least-squares fits of the experimental $\Delta I(t)$ values to equation (2) yield values for $k_{ex,a}$ and $k_{ex,b}$ as functions of temperature (Fig. 4a). The temperature dependence of both $k_{ex,a}$ and $k_{ex,b}$ follows the expected form:

$$k_{ex} = \frac{k_B T}{h} \exp\left(\frac{\Delta S^\ddagger}{R} - \frac{\Delta H^\ddagger}{RT}\right) \quad (3)$$

where $\Delta H^\ddagger$ and $\Delta S^\ddagger$ are, respectively, the apparent activation enthalpy and entropy for chemical exchange with solvent, T is the absolute temperature, h is Planck's constant, and k_B Boltzmann's constant. Extrapolation from the small NMR-accessible temperature range to 298 K according to equation (3) gives rate coefficients

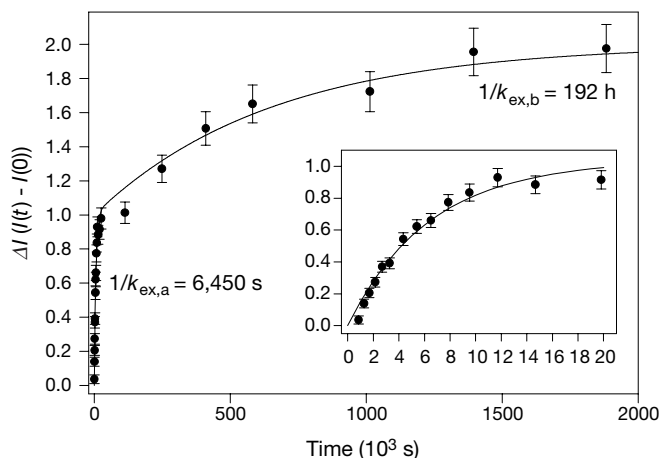


Figure 3 Increase of ^{17}O -NMR intensity of Al_{13} solutions with time after addition of H_2O enriched to 30% in ^{17}O . Data were obtained at $T = 282$ K. Bi-exponential behaviour in near-integral steps indicates differing rates of exchange for the two distinct μ_2 -OH. Inset shows early part of curve at expanded scale. Lines are a least-squares fit to equation (2). Uncertainties were estimated by propagating an uncertainty in the raw relative intensities of 0.02 absolute (1σ ; total intensity normalized to unity). This value was judged to comfortably span the range of values consistent with reasonable adjustments in spectrum phasing and baseline correction.

$k_{\text{ex,a},298} = (1.7 \pm 0.2) \times 10^{-2} \text{ s}^{-1}$ and $k_{\text{ex,b},298} = 1.5 (\pm 0.1) \times 10^{-5} \text{ s}^{-1}$ at pH 5.1–5.4. The associated activation parameters for each site (a, b) were determined by transmitting the uncertainties through a Monte Carlo error analysis and gave values of $\Delta H_a^\ddagger = 211 (\pm 17) \text{ kJ mol}^{-1}$ and $\Delta S_a^\ddagger = 421 (\pm 58) \text{ J mol}^{-1} \text{ K}^{-1}$ for the more labile site (a), and $\Delta H_b^\ddagger = 100 (\pm 9) \text{ kJ mol}^{-1}$, and $\Delta S_b^\ddagger = -10 (\pm 30) \text{ J mol}^{-1} \text{ K}^{-1}$ for the less labile site (b). At 285 K, values of $k_{\text{ex,a}}$ and $k_{\text{ex,b}}$ did not vary measurably over the pH range 4.6 to 5.3 (W.H.C. *et al.*, manuscript in preparation). The activation parameters for the less labile hydroxyl are similar to those reported previously for uncatalysed dissociation of hydroxybridged dimers^{7,8}, whereas those for the more reactive site are much larger than measured for elementary reactions, and probably represent a more complicated reaction pathway.

The exchange rate of the $\eta\text{-OH}_2$ was measured from the variation in the width of the peak at +20 p.p.m. with temperature for isotopically equilibrated samples (Fig. 4b), which decreases with temperature through a minimum near 285 K, then increases. A least-squares fit to the standard rate laws^{8,13,20} gives $\Delta H^\ddagger = 50 \pm 12 \text{ kJ mol}^{-1}$ and $\Delta S^\ddagger = -18 \pm 39 \text{ J mol}^{-1} \text{ K}^{-1}$, yielding a rate coefficient $k_{\text{ex,w},298} = 1,100 \pm 200 \text{ s}^{-1}$ for chemical exchange of the $\eta\text{-OH}_2$ sites. This result is consistent with the immediate appearance of the peak at +20 p.p.m. in the NMR spectra upon addition of labelled water.

These results document an extraordinary range of labilities ($>10^7$) for the different exposed oxy-groups on the model 'Al₁₃' complex, as has been observed before in some molybdate polyoxocations⁸. Furthermore, when compared to data for the dissolution rate of the complex, our data indicate that hydroxyl exchange occurs many times before any irreversible structural change that causes complete dissociation. The time constant for decomposition of the 'Al₁₃' is $t_{1/2} \approx 500 \text{ h}$ under conditions similar

to those of this study¹⁶. This lifetime is much larger than the values $t_{1/2,\text{a},298} = 41 \text{ s}$ and $t_{1/2,\text{b},298} = 13 \text{ h}$. However, the exchange rate coefficient for the terminal water molecules of 'Al₁₃' lies within the range of values observed for monomeric aqueous complexes of Al^{3+} , which vary from 3 s^{-1} for $\text{Al}(\text{H}_2\text{O})_6^{3+}$ (ref. 13) to $30,000 \text{ s}^{-1}$ for $\text{AlOH}(\text{H}_2\text{O})_5^{2+}$ (ref. 17). The rates of exchange of coordinated waters from monomolecular steps on structurally similar surfaces are likely to fall in the same range, although some variation with the charge and solvent-ordering properties of the interface is expected. Solvent-elimination reactions are the rate-controlling step in many macroscopic surface processes, such as adsorption reactions (see, for example, ref. 4), and the 'Al₁₃' results suggest that the data for monomers provide a reasonable range for the rates of these reactions.

The present data raise a number of issues that warrant further study, foremost of which is the large difference in rate and activation enthalpy for the two types of $\mu_2\text{-OH}$. These hydroxyls have identical first-neighbour coordination (two Al and one H), neutral formal charge, and nearly identical Al–O bond distances, based on the crystal structures of the 'Al₁₃' selenate and sulphate salts^{9,21}. Thus, the structural chemistry beyond the first coordination sphere, or interaction with solvent, must greatly influence the exchange process. This result has important implications for modelling surface reactions, since gas-phase *ab initio* calculations commonly used to interpret aqueous surface reactions (see, for example, refs 22–24) use small atomic clusters and do not include the second-sphere and solvent interactions. Qualitative reasoning based on the formal charge distribution and steric considerations do not suggest an unambiguous assignment of rate parameters to $\mu_2\text{-OH}'$ and $\mu_2\text{-OH}$; resolution of this matter awaits further study. The independence of the measured exchange rates with pH indicates that these reactions are not proton-catalysed, whereas complete decomposition of the complex is approximately first-order in protons^{11,16}. This observation, and the apparent lack of isotopic equilibration of the internal $\mu_4\text{-oxo}$, suggest that dissolution of the complex, and perhaps of minerals as well, depends on hydrolysis of highly coordinated oxygens internal to the solid/fluid interface.

The idea that the 'Al₁₃' complex can serve as a model for some surfaces is well accepted¹¹, in part because it condenses intact into a gel that ultimately crystallizes into aluminium (hydr)oxide minerals^{25,26}. The data we report here, however, show that the rates of hydroxyl exchange are extremely sensitive to subtle differences in local structure and Brønsted acidity of the surface. These reactions are elementary, or nearly so, so we expect rates of exchange of oxo groups on model complexes such as 'Al₁₃' to be an essential bridge between the computational simulation of microscopic processes of mass exchange in molecules and surfaces, and the macroscopic phenomena of adsorption and dissolution. □

Received 6 July 1999; accepted 4 February 2000.

- Casey, W. H. & Westrich, H. R. Control of dissolution rates of orthosilicate minerals by divalent metal-oxygen bonds. *Nature* **355**, 147–149 (1992).
- Casey, W. H. & Ludwig, C. The mechanism of dissolution of oxide minerals. *Nature* **381**, 506–509 (1996).
- Stumm, W. *Chemistry of the Solid-Water Interface* (Wiley, New York, 1992).
- Hachiya, K., Sasaki, M., Ikeda, T., Mikami, N. & Yasunaga, T. Static and kinetic studies of adsorption-desorption of metal ions on a $\gamma\text{-Al}_2\text{O}_3$ surface 2. Kinetic study by means of pressure-jump technique. *J. Phys. Chem.* **88**, 27–31 (1984).
- Karlsson, H. R. & Clayton, R. N. Oxygen isotope fractionation between analcime and water: an experimental study. *Geochim. Cosmochim. Acta* **54**, 1359–1368 (1990).
- Lambert, J.-F. & Che, M. in *Dynamics of Surfaces and Reaction Kinetics in Heterogeneous Catalysis* (eds Froment, G. F. & Waugh, K. C.) 91–110 (Elsevier, Amsterdam, 1997).
- Springborg, J. Hydroxo-bridged complexes of chromium(III), cobalt(III), rhodium(III) and iridium(III). *Adv. Inorg. Chem.* **32**, 55–169 (1988).
- Richens, D. T. *The Chemistry of Aqua Ions* (Wiley, Chichester, 1997).
- Johansson, G. The crystal structure of a basic aluminium selenate. *Ark. Kemi* **20**, 305–319 (1962).
- Furrer, G., Ludwig, C. & Schindler, P. W. On the chemistry of the Keggin Al₁₃ polymer. *J. Colloid Interface Sci.* **149**, 56–67 (1992).
- Wehrli, B., Wieland, E. & Furrer, G. Chemical mechanisms in the dissolution kinetics of minerals; the aspect of active sites. *Aquat. Sci.* **52**, 3–31 (1990).
- Hiemstra, T., Yong H. & Van Riemsdijk, W. H. Interfacial charging phenomena of aluminum (hydr)oxides. *Langmuir* **15**, 8045–8051 (2000).
- Hugi-Cleary, D., Helm, L. & Merbach, A. E. Variable temperature and variable pressure ¹⁷O NMR study of water exchange of hexaquaaluminum(III). *Helv. Chim. Acta* **68**, 545–554 (1985).

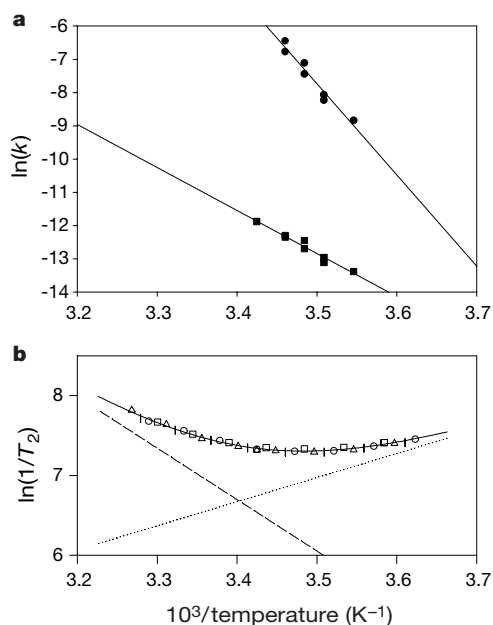


Figure 4 Temperature dependence of oxygen exchange rates for distinct sites on the 'Al₁₃' complex. **a**, Arrhenius plot of the hydroxyl exchange rates (k) with inverse temperature. Least-squares fits to equation (3) yield extrapolated rate coefficients $k_{\text{ex,a},298} = (1.7 \pm 0.2) \times 10^{-2} \text{ s}^{-1}$ and $k_{\text{ex,b},298} = (1.5 \pm 0.1) \times 10^{-5} \text{ s}^{-1}$. Uncertainties (1σ) were estimated by propagating those for the raw intensities (Fig. 3) and $\pm 0.5 \text{ K}$ in temperature. Temperatures were calibrated by the insertion method. **b**, ¹⁷O transverse relaxation rates ($1/T_2 = \pi$ [full-width at half-maximum]) for the $\eta\text{-OH}_2$ of the 'Al₁₃' complex (peak near +20 p.p.m.). Different symbols correspond to distinct sample solutions. Solid line is a least-squares fit to a sum of terms for quadrupolar relaxation (dotted line) and chemical exchange with solvent (dashed line), yielding a chemical exchange rate coefficient $k_{298} = 1,100 \pm 200 \text{ s}^{-1}$. Estimated uncertainty was obtained by assuming a 10% uncertainty in the fitted linewidth.

14. Akitt, J. W. & Elders, J. M. Multinuclear magnetic resonance studies of the hydrolysis of aluminium(III). Part 8. Base hydrolysis monitored at very high magnetic field. *J. Chem. Soc. Dalton Trans.* 1347–1355 (1988).
15. Sullivan, D. J., Nordin, J. P., Phillips, B. L. & Casey, W. H. The rates of water exchange in Al(III)-salicylate and Al(III)-sulfosalicylate complexes. *Geochim. Cosmochim. Acta* **63**, 1471–1480 (1999).
16. Furrer, G., Gfeller, M. & Wehrli, B. On the chemistry of the Keggin Al_{13} polymer – kinetics of proton-promoted decomposition. *Geochim. Cosmochim. Acta* **63**, 3069–3076 (2000).
17. Nordin, J. P., Sullivan, D. J., Phillips, B. L. & Casey, W. H. An ^{17}O NMR study of the exchange of water on $\text{AlOH}(\text{H}_2\text{O})^{2+}(\text{aq})$. *Inorg. Chem.* **37**, 4760–4763 (1998).
18. Thompson, A. R., Kunwar, A. C., Gutowsky, H. S. & Oldfield, E. Oxygen-17 and aluminum-27 nuclear magnetic resonance spectroscopic investigations of aluminum(III) hydrolysis products. *J. Chem. Soc. Dalton Trans.* 2317–2322 (1987).
19. Helm, L., Elding, L. I. & Merbach, A. E. Activation parameters and mechanism for water exchange of tetraaquaplatinum(II) studied by high-pressure oxygen-17 NMR spectroscopy. *Inorg. Chem.* **24**, 1719–1721 (1985).
20. Merbach, A. E. Use of high pressure kinetic studies in determining inorganic substitution mechanisms. *Pure Appl. Chem.* **59**, 161–172 (1987).
21. Parker, W. O'N. Jr., Millini, R. & Kiricsi, I. Metal substitution in Keggin-type tridecameric aluminum-oxo-hydroxo clusters. *Inorg. Chem.* **36**, 571–575 (1997).
22. Lee, M. A., Winter, N. W. & Casey, W. H. Investigation of the ligand exchange reaction for the aqueous Be^{2+} ion. *J. Phys. Chem.* **98**, 8641–8647 (1994).
23. Xiao, Y. T. & Lasaga, A. C. Ab initio quantum mechanical studies of the kinetics and mechanisms of silicate dissolution- $\text{H}^+(\text{H}_2\text{O})^+$ catalysis. *Geochim. Cosmochim. Acta* **58**, 379–5400 (1994).
24. Kubicki, J. D., Blake, G. A. & Apitz, S. E. Molecular orbital calculations for modeling acetate-aluminosilicate adsorption and dissolution reactions. *Geochim. Cosmochim. Acta* **61**, 1031–1046 (1997).
25. Bradley, S. M., Kydd, R. A. & Howe, R. F. The structure of Al-gels formed through base hydrolysis of Al^{3+} aqueous solutions. *J. Colloid Interface Sci.* **159**, 405–412 (1993).
26. Fu, G., Nazar, L. F. & Bain, A. D. Aging processes of alumina sol-gels: Characterization of new aluminum polyoxocations by ^{27}Al NMR spectroscopy. *Chem. Mater.* **3**, 602–610 (1991).

Acknowledgements

We thank J. Nordin, S. Nordin, S. Crawford and D. Sullivan for help in the early stages of this work. This work was supported by the US NSF and the W.M. Keck Solid State NMR Laboratory at UC Davis.

Correspondence and requests for materials should be addressed to W.H.C. (e-mail: whcasey@ucdavis.edu).

Evidence that humans evolved from a knuckle-walking ancestor

Brian G. Richmond & David S. Strait

Department of Anthropology, The George Washington University, 2110 G Street, NW, Washington, DC 20052, USA

Bipedalism has traditionally been regarded as the fundamental adaptation that sets hominids apart from other primates. Fossil evidence demonstrates that by 4.1 million years ago¹, and perhaps earlier², hominids exhibited adaptations to bipedal walking. At present, however, the fossil record offers little information about the origin of bipedalism, and despite nearly a century of research on existing fossils and comparative anatomy, there is still no consensus concerning the mode of locomotion that preceded bipedalism^{3–10}. Here we present evidence that fossils attributed to *Australopithecus anamensis* (KNM-ER 20419)¹¹ and *A. afarensis* (AL 288-1)¹² retain specialized wrist morphology associated with knuckle-walking. This distal radial morphology differs from that of later hominids and non-knuckle-walking anthropoid primates, suggesting that knuckle-walking is a derived feature of the African ape and human clade. This removes key morphological evidence for a *Pan*–*Gorilla* clade, and suggests that bipedal hominids evolved from a knuckle-walking ancestor that was already partly terrestrial.

During knuckle-walking, chimpanzees and gorillas flex the tips of their fingers and bear their weight on the dorsal surface of their middle phalanges (Fig. 1), permitting them to use their hands for terrestrial locomotion while retaining long fingers for climbing

trees. Cineradiographic experiments show that during the middle and late propulsive phases of knuckle-walking, the carpus and metacarpus are slightly extended relative to the radius, and the proximal carpal joint remains static throughout the propulsive movement¹³. These observations indicate that the wrist is held in a stable, locked position during the support phase of knuckle-walking.

The mechanism by which the wrist locks into place consists of several components. The distal radius of African apes exhibits a distally projecting dorsal margin that gives the joint surface a volar tilt (Fig. 1)^{13,14}. Along this dorsal margin, a distally facing notch accommodates a corresponding concavity on the dorsal surface of the scaphoid (Fig. 1). During wrist extension, the spherical, convex portion of the scaphoid's proximal articular surface rotates until the concave portion engages the notch in a close-packed position¹³ (Fig. 1). This locking mechanism stabilizes the radiocarpal joint by limiting wrist extension during the support phase of knuckle-walking. This mechanism is corroborated by data on passive manipulation in which wrist extension is significantly more restricted in *Pan* and *Gorilla* than in *Pongo* and *Hylobates*^{13,15}. Furthermore, chimpanzees do not consistently recruit the major flexor muscles of the wrist during the support phase of knuckle-walking, suggesting that a passive mechanism exists for limiting extension¹⁶.

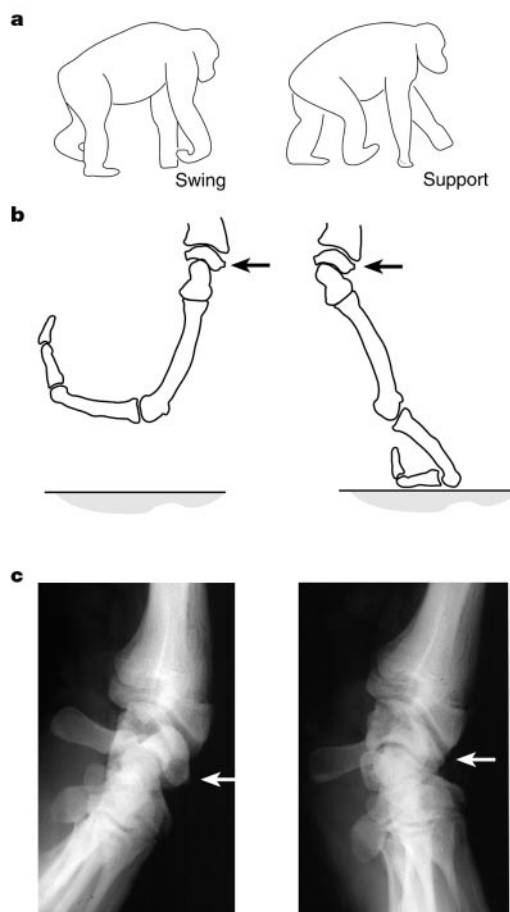


Figure 1 The wrist joint during the swing phase (left column) and support phase (right column) of knuckle-walking. Arrows in **b** and **c** denote the dorsal tip of the scaphoid. **a**, Outline of a chimpanzee during knuckle-walking. **b**, Schematic parasagittal slice of the hand of *Pan troglodytes* along the third ray. The slight concavity on the scaphoid accommodates a convexity on the distal radius. This convexity, the scaphoid notch, lies along the margin of a distally projecting dorsal ridge on the radius. **c**, Radiograph in mediolateral view of the wrist of *P. troglodytes*. Note the close-packed position in the extended wrist (right).

Wrist joint function during knuckle-walking differs from that of other modes of anthropoid pronograde quadrupedalism. Above-branch quadrupedalism in non-hominoid anthropoids typically involves more extended wrist postures along with palmigrady and less extended limbs^{17,18}. Two of the most adept anthropoid terrestrial quadrupeds differ from one another and from knuckle-walkers. Patas monkeys extend their wrists considerably during the late support phase and push-off of terrestrial locomotion¹⁸. In contrast, baboons maintain their wrists near the neutral position when locomoting on the ground, but employ more extended wrist postures while travelling on a branch¹⁷. Thus, wrist function during knuckle-walking seems to be unique compared with other forms of anthropoid quadrupedalism.

Building on published experimental studies^{13,14}, we identify four skeletal features of the distal radius as components of a morphological complex involved in stabilizing the wrist in extension during knuckle-walking. They are (1) the distal projection of the dorsal ridge, (2) the orientation of the scaphoid notch, (3) the size of the scaphoid notch, and (4) the relative orientations of the lunate and scaphoid articular surfaces (Fig. 2a). The distal projection of the dorsal ridge limits extension and might resist shearing forces, and the dorsal position of the notch for the scaphoid ensures that the

radius close-packs with the scaphoid in extension. A relatively large scaphoid notch reduces stress during weight support by increasing the area over which joint forces are distributed. In the radio-ulnar plane, the scaphoid and lunate articular surfaces of the African apes and non-hominoid anthropoids are roughly co-planar and face distomedially, presumably to resist stresses in that plane. In contrast, these surfaces are angled relative to each other in Asian apes, and might contribute to a ball-and-socket mechanism¹³.

A canonical variates analysis (Fig. 2b) based on the distal projection index, scaphoid notch angle, notch size index and scaphoid–lunate facet angle demonstrates that all extant taxa differ significantly from one another ($P < 0.001$ for all pairwise contrasts except *Papio*–*Erythrocebus*, $P < 0.015$). Notch size index and interfacet angle are the most influential variables along canonical axis 1, and the distal projection index dominates canonical axis 2. The Miocene hominoid *Proconsul heseloni* resembles extant howler monkeys, suggesting that the primitive hominoid condition is similar to that of a generalized anthropoid. The knuckle-walking African apes (*Pan* and *Gorilla*) are distinguished in the analysis by a combination of features, including a very prominent distally projecting dorsal ridge, a more dorsally oriented scaphoid notch, and a scaphoid notch and scaphoid–lunate angle that are small compared with those of other anthropoid quadrupeds but larger than those of Asian apes (Fig. 3). A UPGMA clustering diagram of the multivariate Mahalanobis D^2 distances illustrates the similarity between the radii of *A. anamensis* and *A. afarensis* and those of the knuckle-walking African apes (Fig. 2c), indicating that these hominids retain the derived wrist morphology of knuckle-walkers. Thus, the distally projecting dorsal ridge (Fig. 3) and distinct, dorsolaterally oriented scaphoid notch found in radii attributed to *A. anamensis* and *A. afarensis* indicate that wrist extension was limited in these hominids as in knuckle-walkers. The radius attributed to *Paranthropus robustus* is most similar to humans but is somewhat intermediate, and the radius of *A. africanus* is remarkably human-like (Figs 2b, 2c, 3). The sampling limitations of the fossil record do not allow us to assess the significance of variation between hominid species (except perhaps for the substantial difference between *A. africanus* and *A. anamensis/A. afarensis*). However, the hominids sort temporally such that the earliest fossils are the most African ape-like, and the later fossils more closely resemble modern humans, a pattern consistent with the hypothesis presented here.

Previous studies¹⁹ have noted that early hominids lack other distinctive knuckle-walking features, such as pronounced dorsal metacarpal ridges, and dorsally expanded and widened articular surfaces of the metacarpal heads²⁰, that increase the stability and range of extension in metacarpophalangeal joints. However, because these features are not always present in extant knuckle-walkers²⁰, their absence in hominids does not rule out knuckle-walking ancestry. Rather, the absence of these features in early hominids, in conjunction with clearly derived morphological evidence for bipedalism^{1,12,19}, suggests to us that early hominids did not themselves practise knuckle-walking. The total morphological pattern is consistent with the suggestion that the knuckle-walking morphology in the wrists of *A. anamensis* and *A. afarensis* was retained from a knuckle-walking ancestor. These results support earlier interpretations^{5,10} that fusion of the os centrale with the scaphoid, one of the few derived morphological characteristics shared by African apes and humans, is an adaptation to resist weight-bearing stresses and improve the stability of the hand in terrestrial quadrupedalism. Other traits shared between African apes and humans, such as robust metacarpals and phalanges and long middle phalanges relative to proximal phalanges, might also be retained knuckle-walking adaptations²¹.

Our results indicate that the behavioural/morphological complex associated with knuckle-walking is a synapomorphy of the African ape and human clade. This removes key morphological evidence uniting chimpanzees with gorillas to the exclusion of humans²², and

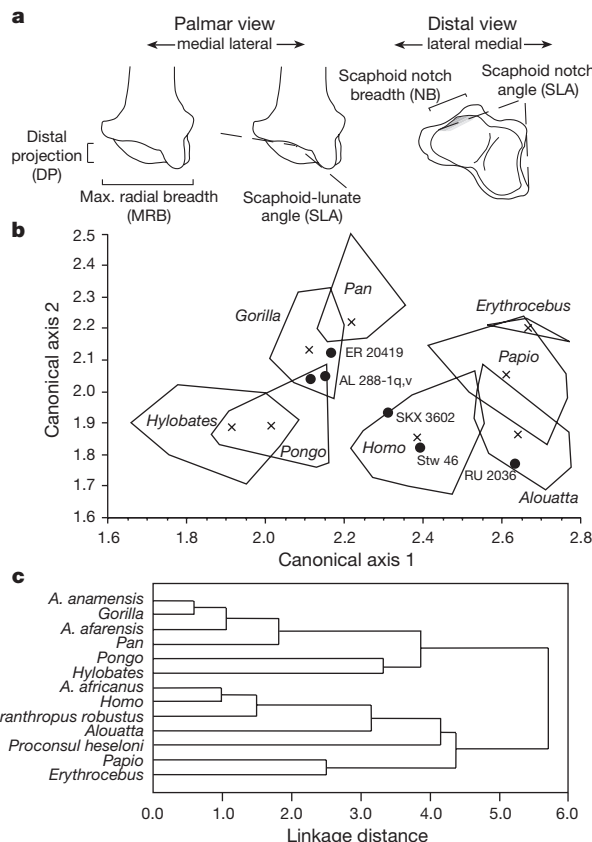


Figure 2 Measurements and canonical variates analysis (CVA) of the distal radius. **a**, Illustration of the measurements collected in palmar and distal views. **b**, Bivariate plot of canonical scores, based on CVA of four shape measurements (DPI, NSI, SNA and SLA). Observed ranges of extant taxa (polygons), taxon means (crosses) and fossil specimens (circles) are shown. The two measures for AL 288-1 come from the left and right radii of this individual. **c**, UPGMA clustering diagram of the Mahalanobis D^2 distances among groups in multidimensional space (not limited to the axes shown in **b**). Note in **b** and **c** that the earliest hominid taxa *A. anamensis* (KNM-ER 20419) and *A. afarensis* (AL 288-1) exhibit the derived knuckle-walking morphology, whereas *Paranthropus robustus* (SKX 3602) and especially *A. africanus* (Stw 46) more closely resemble modern humans.

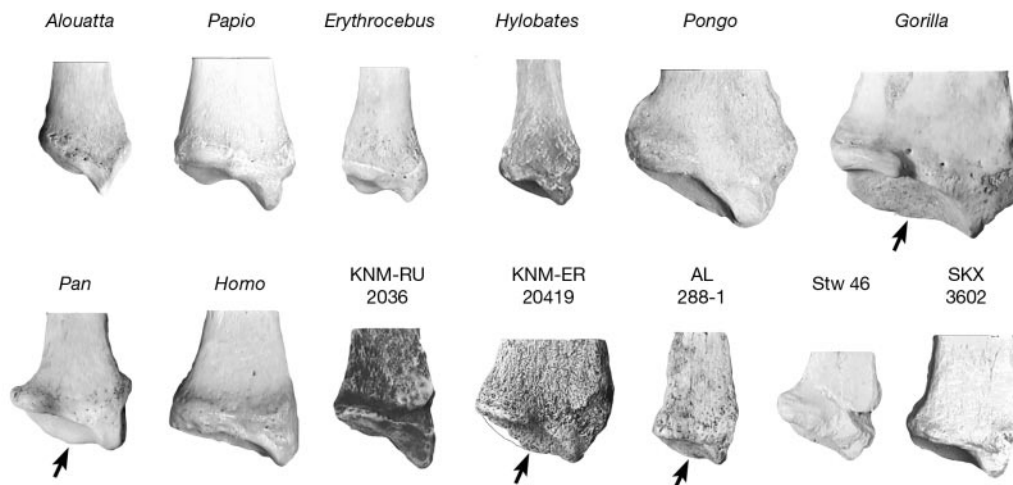


Figure 3 Palmar view of distal radii. All specimens except SKX 3602 (reversed) are left radii. The early hominid radii KNM-ER 20419 (*A. anamensis*) and AL 288-1 (*A. afarensis*) possess the distally projecting and medially extended dorsal ridge (arrows), and intermediate scaphoid notch size and scaphoid–lunate facet angle characteristic of the knuckle-walking African apes. In some specimens (for example, *Papio* and *Pan*), the

scaphoid notch is partly visible in palmar view as a concavity along the dorsal radial margin, just to the left of the styloid process. Note that the dorsal ridge in KNM-ER 20419 had been damaged along its medial extent and was reconstructed here for illustrative purposes. No measurements involved the reconstructed portion.

is consistent with other phylogenies based on morphological²³ and molecular²⁴ data that support a *Pan*–*Homo* clade. This further indicates that knuckle-walking did not originate independently in *Pan* and *Gorilla*. Thus, the differences observed between *Pan* and *Gorilla* in the growth of the ulnar side of the carpus²⁵ might be consequences of kinematic differences related to body size.

The retention of knuckle-walking morphology in the earliest hominids indicates that bipedalism evolved from an ancestor already adapted for terrestrial locomotion. However, early australopithecines also exhibit a number of postcranial traits related to arboreal locomotion, including a long radial neck, well-developed brachioradialis insertion¹¹, superiorly directed glenoid fossa, relatively long forelimbs and relatively long, curved fingers and toes¹⁹. Although there is some debate about the extent to which hominids such as *A. afarensis* continued to practise arboreality^{10,18,19}, the retention of arboreal morphology in early hominids suggests that at least their ancestors engaged in arboreal behaviours. Thus, the wealth of evidence illustrating the biomechanical similarities between climbing and bipedalism⁸ might still help to explain the morphological transition of an ape-like ancestor to an early biped. However, the present study does not support hypotheses in which the ancestor was completely committed to orthograde arboreality^{3,6,7}.

Pre-bipedal locomotion is probably best characterized as a repertoire consisting of terrestrial knuckle-walking, arboreal climbing and occasional suspensory activities⁹, not unlike that observed in chimpanzees today. This raises the question of why bipedalism would evolve from an ancestor already adapted to terrestrial locomotion, and is consistent with models relating the evolution of bipedalism to a change in feeding strategies²⁶ and novel non-locomotor uses of the hands^{4,9,27,28}. □

Methods

Measurements were collected on distal radii attributed to *Proconsul heseloni* (KNM-RU 2036)²⁹, *Australopithecus anamensis* (KNM-ER 20419)¹¹, *Australopithecus afarensis* (AL 288-1q,v)¹², *Australopithecus africanus* (Stw 46)³⁰ and *Paranthropus robustus* (SKX 3602)³⁰ to those of 19 *Hylobates* specimens (pooled *H. syndactylus*, *H. klossi*, *H. lar* and *H. concolor*), 13 *Pan troglodytes*, 17 *Gorilla gorilla*, 17 *Pongo pygmaeus*, 20 *Homo sapiens*, 19 *Papio anubis*, 3 *Erythrocebus patas* and 18 *Alouatta palliata*.

The distal projection (DP) of the dorsal ridge, maximum radial breadth (MRB) and scaphoid–lunate facet angle (SLA) were measured in palmar view, which was defined as the plane of the flattened anterior surface on the distal radial shaft just proximal to the epiphysis. In all anthropoids examined here this plane is roughly perpendicular to the

plane of the ulnar notch, which is oriented parasagittally during fully pronated and fully supinated hand postures, and approximates the plane in which the hand moves during wrist extension. Measurements were collected by digitizing high-resolution video images. In palmar view, each radius was positioned so that the midpoint of the palmar margin of the articular surface lay in the centre of the perpendicularly oriented video camera, and rotated until the lateral shaft margin lay superoinferiorly. Scaphoid notch breadth (NB) was defined as the length of the portion along the dorsal margin of the distal radial articular surface that is bevelled (indicated by shading in Fig. 2a) to articulate with the scaphoid. The scaphoid notch angle (SNA) was measured as the angle between the scaphoid and ulnar notches. The notch size index (NSI) was calculated as NB/MRB, and the distal projection index (DPI) as DP/MRB.

Received 10 August 1999; accepted 10 January 2000.

- Leakey, M. G., Fiebel, C. S., McDougall, I. & Walker, A. C. New four-million-year-old hominid species from Kanapoi and Allia Bay, Kenya. *Nature* **376**, 565–571 (1995).
- White, T. D., Suwa, G. & Asfaw, B. *Australopithecus ramidus*, a new species of early hominid from Aramis, Ethiopia. *Nature* **371**, 306–312 (1994).
- Keith, A. Man's posture: its evolution and disorders. *Br. Med. J.* **1**, 451–454, 499–502, 545–548, 587–590, 624–626, 669–672 (1923).
- Washburn, S. L. Behavior and the origin of man. *Proc. R. Anthropol. Inst. Gr. Br. Ireland* **3**, 21–27 (1967).
- Marzke, M. W. Origin of the human hand. *Am. J. Phys. Anthropol.* **34**, 61–84 (1971).
- Tuttle, R. H. Darwin's apes, dental apes, and the descent of man: normal science in evolutionary anthropology. *Curr. Anthropol.* **15**, 389–398 (1974).
- Stern, J. T. Jr Before bipedalism. *Yrbk Phys. Anthropol.* **19**, 59–68 (1975).
- Fleagle, J. G. et al. Climbing: a biomechanical link with brachiation and bipedalism. *Symp. Zool. Soc. Lond.* **48**, 359–373 (1981).
- Rose, M. D. in *Origine(s) de la Bipédie Chez les Hominidés* (eds Coppens, Y. & Senut, B.) 37–48 (Centre National de la Recherche Scientifique, Paris, 1991).
- Gebo, D. L. Climbing, brachiation, and terrestrial quadrupedalism: historical precursors of hominid bipedalism. *Am. J. Phys. Anthropol.* **101**, 55–92 (1996).
- Heinrich, R. E., Rose, M. D., Leakey, R. E. & Walker, A. C. Hominid radius from the Middle Pliocene of Lake Turkana, Kenya. *Am. J. Phys. Anthropol.* **92**, 139–148 (1993).
- Johanson, D. C. et al. Morphology of the Pliocene partial hominid skeleton (AL288-1) from the Hadar Formation, Ethiopia. *Am. J. Phys. Anthropol.* **57**, 403–451 (1982).
- Jenkins, F. A. Jr & Fleagle, J. G. in *Primate Functional Morphology and Evolution* (ed. Tuttle, R. H.) 213–231 (Mouton, The Hague, 1975).
- Tuttle, R. H. Knuckle-walking and the evolution of hominoid hands. *Am. J. Phys. Anthropol.* **26**, 171–206 (1967).
- Tuttle, R. H. Quantitative and functional studies on the hands of the anthropoidea: I. The Hominoidea. *J. Morphol.* **128**, 309–363 (1969).
- Susman, R. L. & Stern, J. T. Jr Telemetered electromyography of flexor digitorum profundus and flexor digitorum superficialis in *Pan troglodytes* and implications for interpretation of the O.H. 7 hand. *Am. J. Phys. Anthropol.* **50**, 565–574 (1979).
- Schmitt, D. Forelimb mechanics as a function of substrate type during quadrupedalism in two anthropoid primates. *J. Hum. Evol.* **26**, 441–457 (1994).
- Richmond, B. G. *Ontogeny and Biomechanics of Phalangeal Form in Primates*. Thesis, State Univ. New York at Stony Brook (1998).
- Stern, J. T. Jr & Susman, R. L. The locomotor anatomy of *Australopithecus afarensis*. *Am. J. Phys. Anthropol.* **60**, 279–317 (1983).
- Susman, R. L. Comparative and functional morphology of hominoid fingers. *Am. J. Phys. Anthropol.* **50**, 215–235 (1979).

21. Begun, D. R. Knuckle-walking ancestors. *Science* **259**, 294 (1993).
22. Andrews, P. J. & Martin, L. B. Cladistic relationships of extant and fossil hominoids. *J. Hum. Evol.* **16**, 101–118 (1987).
23. Begun, D. R. Miocene fossil hominids and the chimp-human clade. *Science* **257**, 1929–1933 (1992).
24. Ruvoilo, M. Molecular phylogeny of the hominoids: inferences from multiple independent DNA sequence data sets. *Mol. Biol. Evol.* **14**, 248–265 (1997).
25. Dainton, M. & Macho, G. A. Did knuckle walking evolve twice? *J. Hum. Evol.* **36**, 171–194 (1999).
26. Hunt, K. D. The postural feeding hypothesis: an ecological model for the origin of bipedalism. *S. Afr. J. Sci.* **9**, 77–90 (1996).
27. Hewes, G. W. Food transport and the origin of hominid bipedalism. *Am. Anthropol.* **63**, 687–710 (1961).
28. Lovejoy, C. O. The origin of man. *Science* **211**, 341–350 (1981).
29. Napier, J. R. & Davis, P. The forelimb skeleton and associated remains of *Proconsul africanus*. *Foss. Mamm. Afr.* **16**, 1–70 (1959).
30. Grine, F. E. & Susman, R. L. Radius of *Paranthropus robustus* from Member 1, Swartkrans Formation, South Africa. *Am. J. Phys. Anthropol.* **84**, 229–248 (1991).

Acknowledgements

We thank L. Aiello, E. Delson, B. Demes, J. Fleagle, F. Grine, W. Jungers, S. Larson, D. Lieberman, O. Pearson, D. Pilbeam, J. Polk, E. Sarmiento, R. Susman and B. Wood for providing valuable comments; F. Grine, L. Gordon, R. Thorington, R. Potts, A. Walker, and C. Ward for providing access to specimens in their care; and The Henry Luce Foundation.

Correspondence and requests for materials should be addressed to B.G.R. (e-mail: brich@gwu.edu)

Predictive accuracy of population viability analysis in conservation biology

Barry W. Brook[†], Julian J. O'Grady^{*}, Andrew P. Chapman^{*}, Mark A. Burgman[‡], H. Resit Akçakaya[§] & Richard Frankham^{*}

^{*} Key Centre for Biodiversity and Bioresources, Department of Biological Sciences, Macquarie University, New South Wales 2109, Australia

[‡] Environmental Science, School of Botany, University of Melbourne, Parkville, Victoria 3052, Australia

[§] Applied Biomathematics, 100 North Country Road, Setauket, New York 11733, USA

[†] Present address: Key Centre for Tropical Wildlife Management, Northern Territory University, Darwin, Northern Territory 0909, Australia

Population viability analysis (PVA) is widely applied in conservation biology to predict extinction risks for threatened species and to compare alternative options for their management^{1–4}. It can also be used as a basis for listing species as endangered under World Conservation Union criteria⁵. However, there is considerable scepticism regarding the predictive accuracy of PVA, mainly because of a lack of validation in real systems^{2,6–8}. Here we conducted a retrospective test of PVA based on 21 long-term ecological studies—the first comprehensive and replicated evaluation of the predictive powers of PVA. Parameters were

estimated from the first half of each data set and the second half was used to evaluate the performance of the model. Contrary to recent criticisms, we found that PVA predictions were surprisingly accurate. The risk of population decline closely matched observed outcomes, there was no significant bias, and population size projections did not differ significantly from reality. Furthermore, the predictions of the five PVA software packages were highly concordant. We conclude that PVA is a valid and sufficiently accurate tool for categorizing and managing endangered species.

PVA is a way to predict the probability of population (or species) extinction, by inputting actual life-history information and projecting it forward using stochastic computer simulation^{1–4}. PVA is commonly used in a comparative way, to evaluate the effectiveness of different management options; because of this it has been argued that PVA predictions do not need to be precise^{2,9,10}. However, there is no clear dichotomy between relative and absolute predictions—as conservation actions entail costs, management decisions are based not only on whether the proposed strategy is sufficient to achieve recovery, but also on whether the likely benefit will justify the expenditure. These considerations require PVA predictions to be quantitatively reliable. Uncertainties surrounding its predictive reliability have led to conclusions drawn from PVA being rejected in the law courts¹¹. It is therefore essential that the predictions of PVA be compared and tested^{12–14}. Here we assess the predictive accuracy of PVA, and determine whether different generic PVA computer packages differ in their predictive capabilities.

Historical data have been used to test and improve the predictions of complex climatic¹⁵, economic¹⁶, geological¹⁷ and ecological¹⁸ models. As PVA models have important stochastic components, the conclusions of past studies based on a single test^{13,19} lack power, and are prone to case-specific peculiarities²⁰. Consequently, a valid test of PVA predictions must incorporate a large number of data sets to obtain representative assessments. We conducted retrospective analyses on 21 wildlife populations—8 avian, 11 mammalian (representing 9 species), 1 reptilian and 1 piscine.

The 21 data sets used were the only long-term studies we identified that presented data of sufficient duration and quality to be suitable for retrospective testing (see Methods). Five of the most commonly applied 'generic' PVA packages (GAPPS, INMAT, RAMAS Metapop, RAMAS Stage and VORTEX) were used. These packages are all suitable for generic, single-population risk assessments. They offer the most realistic prospects for improving PVA as they are subject to wide scrutiny, are repeatedly used and are frequently revised and updated. All have been used in the management and conservation of endangered species. The key features and differences between the five PVA packages are given in the Supplementary Information.

To avoid circularity, the total data available for each population was split. The first half was used to develop and parameterize the models; the latter half was reserved for testing the accuracy of the PVA predictions. To ensure that the two time periods were kept completely separate, no information from the second half of the population history was used in formulating or parameterizing the

Table 1 Comparison of actual versus predicted quasi-extinction risks

	Expected number	GAPPS [*]	INMAT	R META	R STAGE	VORTEX
No. $N(\text{actual}) < N(Q 90\%)$	18.9	17	16	19	17	17
G-test P		0.48	0.07	0.94	0.21	0.21
No. $N(\text{actual}) < N(Q 50\%)$	10.5	11	10	8	11	8
G-test P		0.65	0.83	0.27	0.83	0.27
No. $N(\text{actual}) < N(Q 10\%)$	2.1	3	1	3	4	3
G-test P		0.48	0.38	0.54	0.21	0.54

The number of actual populations that declined below the predicted population size corresponding to quasi-extinction probabilities of 90%, 50% and 10%, for the population viability analysis software packages GAPPS, INMAT, RAMAS Metapop, RAMAS Stage and VORTEX. The sample is based on 21 retrospective studies. None of the PVA packages' predictions differed significantly from the expected outcome based on goodness of fit (G) tests.

^{*} GAPPS crashes at very large population sizes because of memory limitations, so the fish was not modelled with this package (expected numbers adjusted to 18, 10 and 2).

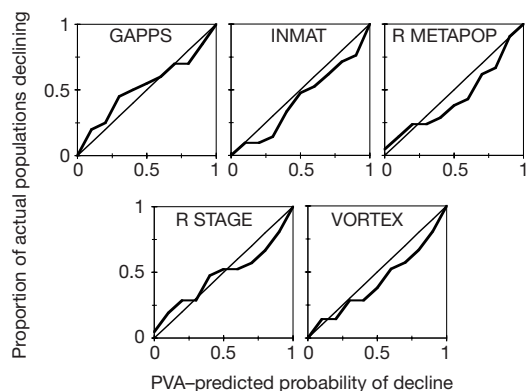


Figure 1 Plot of the PVA-predicted probability of population decline (quasi-extinction risk) versus the actual proportion of the 21 real populations that decline below the corresponding threshold size. These threshold sizes represent different percentage declines in different populations, but are always associated with the same level of risk. For example, half (10.5) of the 21 historical populations should have actually declined below the size assigned a 50% probability by the PVA. For each of the five PVA software packages, a perfect fit with reality lies on the 45° line.

PVA models. All predictive tests were done within species and then pooled across taxa. The accuracy of PVA was assessed by comparing the predicted quasi-extinction risk (the likelihood that the population will decline below a given size in a specified time frame²¹) and population size projections with reality (see Methods).

The results showed a surprisingly close relationship between the PVA predictions and the historical behaviours of the 21 real populations. The quasi-extinction probabilities predicted by the PVAs were not significantly different from the proportion of actual populations that declined below a given threshold (Fig. 1). The risks of decline predicted by the PVAs did not deviate significantly from reality (Table 1), although there was a trend towards slightly pessimistic predictions. Furthermore, there were no overall differences between PVA software packages in their predictions of quasi-extinction risks at 90%, 50% and 10% risk levels (for example, at the 50% quasi-extinction level, Kruskal-Wallis rank test, $H = 1.01$, d.f. = 4, $P = 0.907$).

Actual population sizes consistently fell within the bounds predicted by the stochastic simulations. The projected mean final population sizes were not significantly different from the actual final population sizes (for any of the PVA packages), on the basis of a distribution test combining all 21 studies. Likewise, predicted and actual population growth rates (r) did not differ significantly. There was no consistent over- or underestimation of future population numbers (Fig. 2). A sign test on the predictive bias was not significant for any package, nor when pooled across all packages. There was also no difference in bias across the five PVA packages ($G[\text{Heterogeneity}] = 0.957$, d.f. = 4, $P = 0.916$).

The predictions of the five PVA packages were significantly correlated across the 21 retrospective studies, on the basis of both quasi-extinction probabilities and the fit of projections to future population sizes (Table 2). There were some significant differences

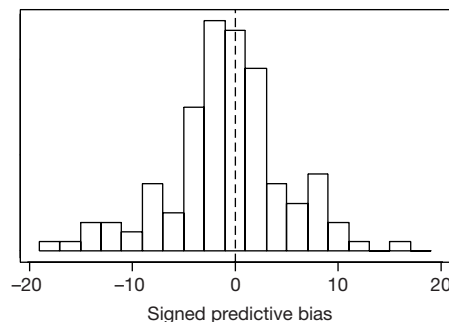


Figure 2 Signed predictive bias in projected population size (compared with actual population numbers), taken across 21 populations for five PVA packages. There is no significant positive or negative bias.

between particular packages for individual species (on the basis of Kolmogorov–Smirnov confidence interval tests²² on quasi-extinction probabilities). However, there was no overall consistency in these differences, and no package gave demonstratively better or worse predictions than any other. RAMAS Metapop and VORTEX gave the best least-squares model fit to actual population size (taken across all 21 species; see Table 3 of Supplementary Information), although a Kruskal-Wallis rank test showed no significant difference among the PVA packages ($H = 7.9$, d.f. = 4, $P = 0.097$).

The short-term predictions of the PVA models were relatively accurate, with a good overall correspondence between simulated and observed outcomes, on the basis of several criteria. The PVA predictions were not overly optimistic, as has been assumed by some¹⁰ (in fact, the general trend was for the PVA models to be slightly pessimistic). Following the precautionary principle, it is preferable to produce circumspect predictions of risk when managing endangered species²³. These findings contradict the widespread view that stochastic population models are poor predictors of a population's future fate. Nevertheless, if the means and/or variances of vital rates were to change substantially in the future, the absolute accuracy of PVA predictions would be questionable. Furthermore, these results may not apply to plants (which were not included in this study), nor to cases where too few life-history or population data are available to estimate model structure and input parameters reliably¹³.

In some cases, there were significant absolute differences between the predictions of different software packages, due to differences in model structure. For example, a ceiling carrying capacity had to be implemented for the cycling Soay and Boreray sheep populations in INMAT, even though this was known to be inappropriate for these populations²⁴. In these cases, the other packages that allowed overcompensatory density-dependent survival to be modelled gave more realistic predictions than INMAT. For a given species, the best packages are likely to be those that most realistically model its life-history. Unfortunately, it is not always possible to discern the type of model that fits a particular case *a priori*. Despite the best efforts of PVA modellers, it is possible that some events (such as catastrophes) may be overlooked, making it impossible to guarantee

Table 2 Correlations between the predictions of five PVA packages

	Probability of decline*				Model fit to actual N_t †			
	INMAT	R META	R STAGE	VORTEX	INMAT	R META	R STAGE	VORTEX
GAPPS	0.83	0.78	0.86	0.86	0.66	0.94	0.88	0.90
INMAT		0.63	0.87	0.76		0.70	0.82	0.78
R META			0.83	0.87			0.94	0.89
R STAGE				0.88				0.89

* Probability of the maximum observed historical population decline.

† Scaled fit of model projections to actual population size (N_t).

All correlations are highly significant ($P < 0.01$).

that incorrect predictions will not be made. It is therefore important to explore alternative model structures and conduct sensitivity analyses, particularly when data are scarce^{2,3,14}. Moreover, information collected from past monitoring should be used to test and refine PVA models, forming a continual feedback process of development and improvement²⁵. PVA is the best tool we have for estimating extinction risk, and the alternatives are subjective, less rigorous, and likely to provide poorer predictions²⁶.

PVA predictions are surprisingly accurate, given adequate data, and should be useful in the conservation contexts in which they are currently applied. There are also high correlations between the predictions of different PVA packages. These results provide strong empirical justification for the use of PVA for categorizing the vulnerability of endangered animal species and evaluating options for their recovery. Furthermore, they validate PVA as a useful research tool for addressing unresolved issues in conservation biology. □

Methods

Data sets

The protocol for choosing examples was independent of their structure, detail and outcome. The criteria used were: (1) a minimum duration of 10 years; (2) data of sufficient quality and detail to build PVA models; (3) a small population size (very large populations are generally not of conservation concern); (4) data concerning endangered species or isolated populations. Thus, the filter was based on the priorities of our colleagues in collecting ecological data. The 21 populations that we were able to use (see Supplementary Information) covered a range of taxa (birds, mammals, reptiles and fish), tropic levels (omnivores, herbivores and carnivores), environmental conditions (low to high environmental variability) and insular and mainland populations, encompassed several different modes of population regulation (exponential growth or decline, a population ceiling, and density dependence), and spanned a range of geographical and climatic zones (Africa, Europe, North America and Oceania). Nevertheless, not all taxa are completely or proportionally represented in our sample, and the focus is on extant species.

PVA model structure

The structure of each PVA model depended on the biology of the given species and the built-in features of each software package (see Supplementary Information). To ensure that the models were as realistic as possible, all relevant aspects of the population's ecology were included (within the capabilities of each PVA package). The exact structure of the PVA models built using different packages often differed for a given species, depending on the available features, limitations and assumptions associated with each program²⁷. This made it possible to test whether some packages were better predictors than others because of their different features.

Parameter estimation

We used the best estimates for each parameter, given the available demographic, environmental and population data. Where the estimated population parameters were given directly in the literature, we checked them independently when possible. In most cases the compilation and analysis of data were done in collaboration with the people who worked on the taxon in question. The protocols used to estimate the PVA model parameters are presented in the Supplementary Information.

PVA runs

Five hundred stochastic simulation runs were performed in each case to ensure statistical reliability.

Testing PVA predictions

The predictive accuracy of PVA was assessed as follows. (1) Quasi-extinction probabilities were used as a surrogate for absolute extinction risk, as only one study population went extinct. If the PVA predictions are correct, the proportion of historical populations that declined below a given threshold should equate with the proportion of simulation trajectories that declined below the same threshold. To test this, the number of actual populations (out of 21) that declined below the threshold population sizes corresponding to predicted quasi-extinction probabilities of 0%, 10%, 20%, ..., 100% were recorded. (2) Mean population size predictions and growth rates were compared with reality using a combined test based on the sum of standard deviates²⁸. The objective was to test for equality over multiple populations, and combine the results of individual paired tests to increase statistical power. A log-transformation was used to normalize population size data. (3) Signed predictive bias (mean error): models with good predictive properties will produce a bias close to zero. Bias was calculated as $\Sigma(\text{Predicted size} - \text{Actual size}) / \text{number of years } (N)$. This was converted to a normal deviate $W = \sqrt{N} \times \text{bias}/s$, where s is the standard deviation of bias values. (4) The 'fit' of predicted versus actual population size over time was used to assess projections of future population numbers, calculated as the sum of squared deviations between observed and actual population sizes $(\Sigma[\text{Predicted} - \text{Actual}]^2 / \text{Actual})$ for each simulation replicate, averaged over 500 repli-

cates. (5) Pearson's product moment correlations were used to determine concordance amongst the predictions of different PVA packages for quasi-extinction risk and population size projections.

Statistical assumptions

This study represents a general 'meta-analysis' problem²⁹, the units of analysis being independent studies rather than individual subjects. The statistical population was composed of all studied biological populations with sufficient data, and was sampled completely. The underlying assumption of all our statistical tests on the combined results was that the 'errors' were independent among the 21 cases²⁷. As the case studies had little overlap in terms of place, time-period or researcher, this is a valid assumption. The quasi-extinction analysis was based on a simple goodness of fit test, carrying no assumption about the distributions of individual cases. For tests comparing different PVA packages, we verified that the assumptions of parametric tests^{13,14} were met.

Received 7 September 1999; accepted 4 February 2000.

- Boyce, M. S. Population viability analysis. *Annu. Rev. Ecol. Syst.* **23**, 481–506 (1992).
- Beissinger, S. R. & Westphal, M. I. On the use of demographic models of population viability in endangered species management. *J. Wildl. Mgmt.* **62**, 821–841 (1998).
- Groom, M. J. & Pascual, M. A. in *Conservation Biology* (eds Fiedler, P. L. & Kareiva, P. M.) 4–27 (Chapman & Hall, New York, 1998).
- Norton, T. W. (ed.) Special issue: applications of population viability analysis to biodiversity conservation. *Biol. Conserv.* **73**, 91–176 (1995).
- International Union for the Conservation of Nature. *IUCN Red List Categories* (Gland, Switzerland, 1994).
- Caughley, G. Directions in conservation biology. *J. Anim. Ecol.* **63**, 215–244 (1994).
- Taylor, B. L. The reliability of using population viability analysis for risk classification of species. *Conserv. Biol.* **9**, 551–558 (1995).
- Ludwig, D. Is it meaningful to estimate a probability of extinction? *Ecology* **80**, 298–310 (1999).
- Possingham, H. P., Lindenmayer, D. B. & Norton, T. W. A framework for the improvement of threatened species based on population viability analysis (PVA). *Pac. Conserv. Biol.* **1**, 38–45 (1993).
- Ralls, K. & Taylor, B. L. in *The Ecological Basis of Conservation: Heterogeneity, Ecosystems, and Biodiversity* (eds Pickett, S. B. A., Ostfeld, R. S., Shachak, M. & Likens, G. E.) 228–235 (Chapman & Hall, New York, 1997).
- Talbot, J. *Land and Environment Court of New South Wales: Record of Hearing #10151 of 1994* (Court Report, Sydney, 1994).
- Soulé, M. E. & Kohm, K. A. *Research Priorities for Conservation Biology* (Island, Washington DC, 1989).
- Brook, B. W., Lim, L., Harden, R. & Frankham, R. Does population viability analysis software predict the behaviour of real populations? A retrospective study on the Lord Howe Island woodhen *Tricholimnas sylvestris* (Slater). *Biol. Conserv.* **82**, 119–128 (1997).
- Brook, B. W., Cannon, J. R., Lacy, R. C., Mirande, C. & Frankham, R. A comparison of the population viability analysis packages GAPPs, INMAT, RAMAS and VORTEX for the Whooping crane (*Grus americana*). *Anim. Conserv.* **2**, 23–31 (1999).
- Wigley, T. M. L. Climate change: a successful prediction? *Nature* **376**, 463–464 (1995).
- Ramanathan, R. *Introductory Econometrics with Applications* 2nd edn (Dryden, Fort Worth, 1995).
- Oreskes, N., Shrader-Fenchette, K. & Belitz, K. Verification, validation, and the confirmation of numerical models in the earth sciences. *Science* **263**, 641–646 (1994).
- Wahlberg, N., Moilanen, A. & Hanski, I. Predicting the occurrence of endangered species in fragmented landscapes. *Science* **273**, 1536–1538 (1996).
- Huggard, D. J. A linear programming model of herbivore foraging: imprecise, yet successful? *Oecologia* **100**, 470–474 (1994).
- Belovsky, G. E. How good must models and data be in ecology? *Oecologia* **100**, 475–480 (1994).
- Ginzburg, L. R., Slobodkin, L. B., Johnson, K. & Bindman, A. G. Quasi-extinction probabilities as a measure of impact on population growth. *Risk Anal.* **2**, 171–182 (1982).
- Sokal, R. R. & Rohlf, F. J. *Biometry* 4th edn (Freeman, New York, 1995).
- Akçakaya, H. R., Burgman, M. A. & Ginzburg, L. R. *Applied Population Ecology* 2nd edn (Sinauer, Sunderland, Massachusetts, 1999).
- Grenfell, B. T., Price, O. F., Albon, S. D. & Clutton-Brock, T. H. Overcompensation and population cycles in an ungulate. *Nature* **355**, 823–826 (1992).
- Burgman, M. A., Ferson, S. & Akçakaya, H. R. *Risk Assessment in Conservation Biology* (Chapman & Hall, London, 1993).
- Zeckhauser, R. J. & Viscusi, W. K. Risk within reason. *Science* **248**, 559–564 (1990).
- Mills, L. S. *et al.* Factors leading to different viability predictions for a grizzly bear data set. *Conserv. Biol.* **10**, 863–873 (1996).
- Reynolds, M. R., Burkhart, H. E. & Daniels, R. F. Procedures for statistical validation of stochastic simulation models. *Forest Sci.* **27**, 349–364 (1981).
- Arnqvist, G. & Wooster, D. Meta-analysis: synthesizing research findings in ecology and evolution. *Trends Ecol. Evol.* **10**, 236–240 (1995).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

Acknowledgements

We thank K. Armitage, M. Boyce, J. Cannon, C. Catterall, T. Clutton-Brock, M. Curry, B. Grenfell, R. Harden, J. Kikkawa, M. Lynch, C. Mirande, R. Peterson, C. Schwartz, P. Smith, J. Vucetich and the Interagency Grizzly Bear Study Team for supplying data, and S. Andelman, J. Ballou, J. Bull, S. Ferson, I. Hanski, R. Harris, J. Kikkawa, R. Lacy, H. McCallum, M. McCarthy, S. Mills, P. Miller, S. Pimm, H. Regan and M. Soulé for comments on the manuscript. This study was funded by Australian Research Council and Macquarie University grants.

Correspondence and requests for materials should be addressed to B.W.B. (e-mail: barry.brook@ntu.edu.au).

Low variability in a Y-linked plant gene and its implications for Y-chromosome evolution

Dmitry A. Filatov*, Françoise Monéger†, Ioan Negrutiu† & Deborah Charlesworth*

* Institute of Cell, Animal, and Population Biology, University of Edinburgh, Edinburgh, EH9 3JT, UK

† Ecole Normale Supérieure de Lyon, UMR 5667 CNRS/INRA/ENS/Lyon I, 46 Allée d'Italie, 69364 Lyon, France

Sex chromosomes have evolved independently in several different groups of organisms, but they share common features, including genetic degeneration of the Y chromosome^{1,2}. Suppression of recombination between ancestral proto-X and proto-Y chromosomes is thought to have led to their gradual divergence, and to degeneration of the Y chromosome², but the evolutionary forces responsible are unknown. In non-recombining Y chromosomes, deleterious mutations may be carried to fixation by linked advantageous mutations ("selective sweeps")³. Occurrence of deleterious mutations may drive "Muller's ratchet" (stochastic loss of chromosomes with the fewest mutations)^{2,4}. Selective elimination of deleterious mutations, causing "background selection"^{5,6} may accelerate stochastic fixation of mildly detrimental mutations². All these processes lower effective population sizes, and therefore reduce variability of genes in evolving Y chromosomes. We have studied DNA diversity and divergence in a recently described X- and Y-linked gene pair⁷ (*SLX-1* and *SLY-1*) of the plant *Silene latifolia* to obtain evidence about the early stages of Y degeneration. Here we show that DNA polymorphism in *SLY-1* is 20-fold lower than in *SLX-1*, but the pattern of polymorphism does not suggest a selective sweep.

White campion *Silene latifolia* (Caryophyllaceae) is a short-lived perennial plant species that is dioecious (individuals are male or female) and has a mammal-like chromosomal sex-determination system^{8,9} (XX female and XY male). The Y chromosome is important in the suppression of female and promotion of male flower characters⁹. *S. latifolia* X and Y chromosomes, like those of mammals, have a small recombining pseudo-autosomal region⁹, estimated to be about 10% of the Y chromosome. The ancestor of *S. latifolia* seems to have diverged from the closest non-dioecious species about 10–20 million years ago¹⁰, and the morphologically distinct sex chromosomes probably evolved more recently. The relatively recent origin of the *S. latifolia* sex chromosomes provides

an opportunity to study the evolutionary forces that operate in the early stages of sex chromosome evolution. Animal sex chromosome systems (such as those in humans¹¹, hominids¹², rodents^{13,14} and *Drosophila*¹⁵) are ancient, and their Y chromosomes are almost completely degenerated. The only data about the early stages of genetic degeneration have been obtained from estimates of DNA polymorphism on a *Drosophila americana* neo-Y, a recent translocation onto the Y chromosome that occurred about one million years ago. These data show only a small decrease in DNA polymorphism relative to the neo-X (ref. 16), which indicates that genetic diversity may be lost slowly from the Y chromosome. Despite its relatively recent origin, the *S. latifolia* Y chromosome shows some signs of degeneration^{8,9}, and plants that have a Y but no X chromosome are usually inviable^{17,18}. The *S. latifolia* Y chromosome is bigger than the X, probably owing to the accumulation of repetitive DNA; several attempts to isolate active Y-linked genes from *S. latifolia* yielded only repetitive sequences¹⁹. The first X-linked gene identified in this species seems to have a degenerated Y-linked homologue²⁰.

Only one active Y-linked gene (*SLY-1*) has so far been characterized in *S. latifolia*⁷. Its non-Y-linked homologue was believed to be located on the X chromosome and was named *SLX-1*, although its X-linkage had not been proved. We tested this by observing the segregation of a molecular genetic marker (an *HpaI* restriction site) in the progeny of parents with the restriction site present (homozygous mother) or absent (father). All ten male progeny inherited the maternal variant, and five females were all heterozygous for the maternal and paternal alleles. We therefore conclude that *SLX-1* is X-linked.

The coding regions of *SLX-1* and *SLY-1* have very similar sequences⁷, which may indicate that they are located in the recombining pseudo-autosomal region of the sex chromosomes. Although segregation tests suggest complete linkage of *SLY-1* to sex⁷, these data are from a family of moderate size, and the probability of detecting rare recombination is low. However, rare recombination can be detected by examining sequence variants. If recombination occurs, some variant sites should be polymorphic in both genes in natural populations, and there should be no fixed differences between the genes. We compared sequences of the same 2-kilobase (kb) region of both *SLX-1* and *SLY-1* (Fig. 1) from 12 males: 3 were from independent laboratory strains, and the remainder were from US ($n = 2$), Scottish ($n = 5$) and Portuguese ($n = 2$) natural populations.

The *SLY-1* sequences showed five segregating nucleotide substitutions (four in introns and a Pro/Leu amino-acid replacement at position 349 in exon 13 of the *SLY-1* protein) and one seven-nucleotide insertion–deletion (indel) polymorphism in intron 13 (Fig. 1). To quantify the diversity values²¹ we estimated two

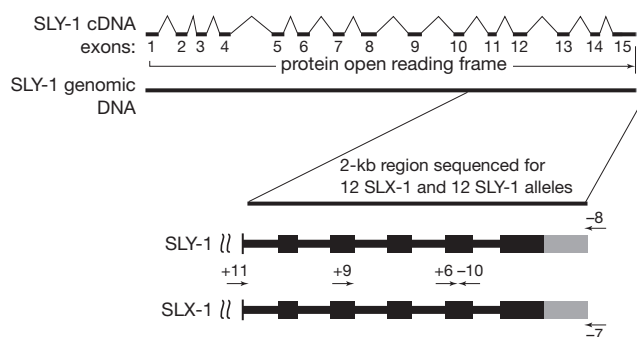


Figure 1 Location of the 2-kb 3' region of the *SLX-1* and *SLY-1* genes sequenced in this study. The entire *SLY-1* gene (C. Delichère, personal communication) is shown, to specify the intron numbers. For the region studied, we determined the intron positions (thin blocks) in both *SLX-1* and *SLY-1*; they are in the same positions in both genes. Exon

numbers are shown below the *SLY-1* cDNA. Non-coding 3' regions of the genes are shown in grey. The positions and directions of primers are shown by arrows. Primers '+11', '+9', '+6' and '+10' are homologous to both *SLX-1* and *SLY-1*. Primers '-7' and '-8' are gene-specific.

commonly used measures of variability, $\theta = 0.00082 \pm 0.00046$, and $\pi = 0.0010 \pm 0.0002$, excluding the indel variant. When the indel was treated as a single variant, $\theta = 0.00098 \pm 0.00053$ and $\pi = 0.00122 \pm 0.00023$. In contrast, the same 2-kb region of the *SLX-1* alleles revealed 95 segregating nucleotide substitutions (5 replacement substitutions and 90 silent variants) and 33 indel polymorphisms (all in introns). Excluding indel regions, $\theta = 0.016 \pm 0.0016$ and $\pi = 0.016 \pm 0.0013$.

Between *SLX-1* and *SLY-1*, there are 54 fixed nucleotide differences and no shared polymorphic sites. The *SLX-1* sequences, but not those of *SLY-1*, show evidence of recombination (for the 12 *SLX-1* allele sequences the "four gamete test"²² detects a minimum of 16 recombination events, and the estimate of recombination per nucleotide²³ is 0.015). The contrast between *SLX-1* and *SLY-1* shows that *SLY-1* is located in the differential segment of the Y chromosome.

The net silent site divergence, K_s (ref. 21), between the *SLX-1* and *SLY-1* sequences is 3%. Assuming a synonymous site molecular clock with a rate of about 0.6% per million years, the best current estimate for plant nuclear genes²⁴, this corresponds to a total divergence of about 5 million years between the X and Y genes, and thus about 2.5 million years since they stopped recombining. This is less than the putative age of the sex chromosomes¹⁰ and much less than the divergence between the MROS3-X and MROS3-Y genes²⁰. This indicates that cessation of recombination may have occurred at different times for different loci, possibly owing to migration of the boundary between the pseudo-autosomal region and non-recombining sections of the Y chromosome.

The apparent recent halting of recombination provides an opportunity to analyse variation in *SLY-1* and to infer the evolutionary forces involved in the degeneration of the Y chromosome. If no selective factors (genetic hitch-hiking, background selection and so on) were operating, the effective population size of a Y-linked gene would be one-quarter of that for an autosomal gene, and one-third of that for an X-linked gene. Thus, in the neutral case, *SLY-1* should have roughly one-third of the DNA variation observed in *SLX-1*. However, *SLY-1* DNA polymorphism is reduced by a factor of 20 compared with that of *SLX-1* (Fig. 2). Neutral coalescent simulations²⁵, based on 32 segregating sites (one-third of the number in *SLX-1*), show that the diversity in *SLY-1* is significantly ($P < 0.0001$) reduced.

What process decreased *SLY-1* DNA diversity so greatly in this short time? Sexual selection can decrease the Y effective population size compared with that of X chromosomal loci. This cannot be excluded, but is unlikely to have had such an extreme and rapid effect in an insect-pollinated plant². The difference in *SLX-1* and *SLY-1* diversity could be due to a bottleneck in effective population size for Y chromosomes. However, this is not consistent with the data because the extent of *SLX-1*/*SLY-1* divergence implies that there was sufficient time for *SLY-1* diversity to recover after recombination ceased. A recent selective sweep at a locus on the *S. latifolia* Y chromosome, eliminating diversity throughout the species, also appears unlikely, because existing *SLY-1* polymorphisms should then be variants accumulated since the selective sweep and should thus mostly be singletons. We found, however, that none of the six *SLY-1* polymorphic sites (including the indel variant) are

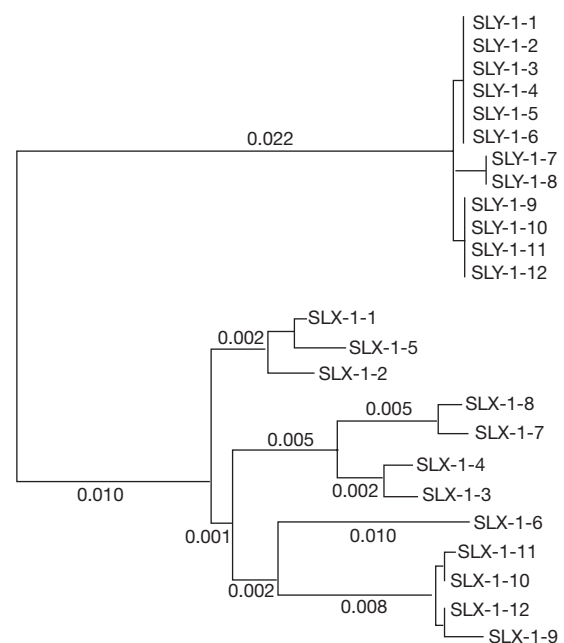


Figure 2 Phylogenetic neighbour-joining tree for a 2-kb region of *SLY-1* and *SLX-1*, constructed using 12 sequences of each of the X- and Y-linked alleles. Jukes–Cantor distances are shown.

singletons, and no significant deviation from neutrality was detected in these data by Tajima's D (ref. 26) and Fu and Li's D^* and F^* statistics²⁷, which were positive, not negative as would be expected after a recent selective sweep (data not shown). Finally, we generated samples of 12 sequences with 32 segregating sites (one-third of the number in the *SLX-1* sample) by coalescent simulations, assuming selective sweeps at different times in the sample's history. The simulations assumed no recombination and used several different strengths of selection. Given the estimated reduction in the number of segregating sites in the Y-linked gene, the number of generations since the selective sweep probably does not exceed the effective population size²¹ of the species; longer times yielded numbers of segregating sites greater than were seen in less than 5% of simulations. Assuming this length of time, or longer, the simulations almost invariably yielded frequencies of single substitutions that were much higher than those observed (Table 1). Thus a simple, classical selective sweep cannot explain the low diversity of the *SLY-1* gene. This differs from the interpretation of low sequence variability in a Y-linked dynein gene in *Drosophila*¹⁵, which may be attributable to selective sweeps, rather than deleterious mutations, because *D. melanogaster* Y chromosomes carry few genes and thus a high deleterious mutation rate seems unlikely. However, our results do not rule out background selection, which does not greatly affect the shape of the gene tree⁵. An explanation involving Muller's ratchet may also be compatible with our data, but no detailed study has yet been done of the ratchet's effects on diversity at neutral

Table 1 Comparison of the observed *SLY-1* sample with simulated samples

Number of generations since selective sweep	Segregating sites			Proportion of single substitutions			Tajima's D^{27}		
	5%	Mean	95%	5%	Mean	95%	5%	Mean	95%
$4N_e$	14	28.02	49	0.095	0.369	0.667	-1.511	-0.082	1.478
$2N_e$	10	19.64	32	0.148	0.463	0.765	-1.522	-0.477	0.893
N_e	6	11.97	21	0.238	0.596	0.900	-1.798	-0.956	0.092
$1/2N_e$	3	7.31	7	0.333	0.722	1.000	-1.944	-1.304	-0.382
Observed		6			0			0.956	

$\theta = 0.005$, due to a selective sweep reducing diversity from an initial level of about one-third of that found in *SLX-1*. Means and percentile values (5% and 95%) are shown for three properties of the simulated samples, for different times since the selective sweep event in units of N_e , the effective population size²¹.

sites in a non-recombining chromosome, and more theoretical work is needed. □

Methods

We used the sequences of *SLX-1* and *SLY-1* to design the following primers for the polymerase chain reaction (PCR) and sequencing (Fig. 1): '–7', 5'-ACTTGCAACG ACTTCACTTTGAG-3'; '–8', 5'-ATCGAATTCAGTGGGAAGTCCA-3'; '+11', 5'-AAGCTCACATGCTGATCTTC-3'; '+9', 5'-GCTGAAGATGGCTTGCTAAC-3'; '+6', 5'-TGGACTTCCACTGGAATTCGAT-3'; '–10', 5'-TCCAGCAGAGCTTGAACAGTC-3'.

A standard cetyltrimethylammonium bromide plant miniprep method with several modifications²⁸ was used to isolate the total genomic DNA from 12 *S. latifolia* males and from 17 plants (2 parents, 10 sons and 5 daughters) used in the *SLX-1* segregation analysis. The 2-kb 3' region in 12 *S. latifolia* males was amplified using the following primer pairs: '+11' and '–7' for *SLX-1*; '+11' and '–8' for *SLY-1*. Because we studied sequences from the X and Y chromosomes of males, direct sequencing was possible for all individuals. Amplification products were passed through 1% agarose gels, column-purified (Qiagen gel extraction kit) and sequenced with the primers listed above, using an ABI Prism 377 automatic sequencer (Perkin Elmer).

The primer pair '+6' and '–7' was used to amplify the short (0.6 kb) 3' *SLX-1* region used in the *SLX-1* segregation analysis. PCR products of both parents were directly sequenced and these sequences were used to choose a restriction enzyme (*HpaI*) to cut the PCR product of only one of the parents. The restriction site was used as a molecular genetic marker in the *SLX-1* segregation analysis.

For the coalescent simulations, we used the program ProSeq v.2.4 (D. Filatov, unpublished, available at <http://helios.bto.ed.ac.uk/evolgen/filatov/proseq.html>). All simulations assumed zero recombination. Simulations to estimate the significance of the lack of diversity in the *SLY-1* gene were conditioned on the observed numbers of segregating sites. Simulations with a selective sweep were conducted according to ref. 29. The neighbour-joining tree was constructed using MEGA software³⁰, based on pairwise divergence values, with Jukes-Cantor correction.

Received 19 August 1999; accepted 28 January 2000.

1. Bull, J. J. *Evolution of Sex Determining Mechanisms*. (Benjamin/Cummings, Menlo Park, 1993).
2. Charlesworth, B. The evolution of chromosomal sex determination and dosage compensation. *Curr. Biol.* **6**, 149–162 (1996).
3. Rice, W. R. Genetic hitchhiking and the evolution of reduced genetic activity of the Y sex chromosome. *Genetics* **116**, 161–167 (1987).
4. Manning, J. T. & Thompson, D. J. Muller's ratchet and the accumulation of favourable mutations. *Acta Biotheoretica* **33**, 219–225 (1984).
5. Charlesworth, D., Charlesworth, B. & Morgan, M. T. The pattern of neutral molecular variation under the background selection model. *Genetics* **141**, 1619–1632 (1995).
6. Orr, H. A. & Kim, Y. An adaptive hypothesis for the evolution of the Y chromosome. *Genetics* **150**, 1693–1698 (1998).
7. Delichère, C. *et al.* *SLY1*, the first active gene cloned from a plant Y chromosome encodes a WD-repeat protein. *EMBO J.* **18**, 4169–4179 (1999).
8. Grant, S. *et al.* Genetics of sex determination in flowering plants. *Dev. Genet.* **15**, 214–230 (1994).
9. Westergaard, M. The mechanism of sex determination in dioecious flowering plants. *Adv. Genet.* **9**, 217–281 (1959).
10. Desfeux, C. *et al.* Evolution of reproductive systems in genus *Silene*. *Proc. R. Soc. Lond. B.* **263**, 409–414 (1996).
11. Anagnostopoulos, T., Green, P. M., Rowley, G., Lewis, C. M. & Giannelli, F. DNA variation in a 5-Mb region of the X-chromosome and estimates of sex-specific/type-specific mutation rates. *Am. J. Hum. Genet.* **64**, 508–517 (1999).
12. Burrows, W. & Ryder, O. A. Y-chromosome variation in great apes. *Nature* **385**, 125–126 (1997).
13. Lundrigan, B. L. & Tucker, P. K. Tracing paternal ancestry in mice, using the Y-linked sex-determining locus *Sry*. *Mol. Biol. Evol.* **11**, 483–492 (1994).
14. Nachman, M. W. & Aquadro, C. F. Polymorphism and divergence at the 5' flanking region of the sex determining locus, *Sry*, in mice. *Mol. Biol. Evol.* **11**, 539–547 (1994).
15. Zurocova, M. & Eanes, W. F. Lack of nucleotide polymorphism in the Y-linked sperm flagellar dynein gene *Dhc-Yh3* of *Drosophila melanogaster* and *D. simulans*. *Genetics* **153**, 1709–1715 (1999).
16. McAllister, B. & Charlesworth, B. Reduced sequence variability on the neo-Y chromosome of *Drosophila americana americana*. *Genetics* **153**, 221–233 (1999).
17. Ye, D. *et al.* Sex determination in the dioecious *Melandrium*. I. First lessons from androgenic haploids. *Sex. Plant Reprod.* **3**, 179–186 (1990).
18. Vagera, J., Paulikova, D. & Dolezel, J. The development of male and female regenerants by *in vitro* Androgenesis in dioecious plant *Melandrium album*. *Ann. Botany* **73**, 455–459 (1994).
19. Buzek, J. *et al.* Isolation and characterization of X chromosome-derived DNA sequences from a dioecious plant *Melandrium album*. *Chromosome Res.* **5**, 57–65 (1997).
20. Guttman, D. S. & Charlesworth, D. An X-linked gene with a degenerate Y-linked homologue in a dioecious plant. *Nature* **393**, 263–265 (1998).
21. Nei, M. *Molecular Evolutionary Genetics*. (Columbia Univ. Press, New York, 1987).
22. Hudson, R. R. & Kaplan, N. L. Statistical properties of the number of recombination events in the history of a sample of DNA sequences. *Genetics* **111**, 147–164 (1985).
23. Hudson, R. R. Estimating the recombination parameter of a finite population model without selection. *Genet. Res.* **50**, 245–250 (1987).
24. Gaut, B. S. Molecular clocks and nucleotide substitution rates in higher plants. *Evol. Biol.* **30**, 93–120 (1998).
25. Hudson, R. R. Gene genealogies and the coalescent process. *Oxf. Surv. Evol. Biol.* **7**, 1–44 (1990).
26. Fu, Y.-X. & Li, W.-H. Statistical tests of neutrality of mutations. *Genetics* **133**, 693–709 (1993).
27. Tajima, F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**, 585–595 (1989).
28. Filatov, D. A. & Charlesworth, D. DNA polymorphism, haplotype structure and balancing selection in the *Leavenworthia* PgiC locus. *Genetics* **153**, 1423–1434 (1999).

29. Braverman, J. M., Hudson, R. R., Kaplan, N. L., Langley, C. H. & Stephan, W. The hitchhiking effect on the site frequency spectrum of DNA polymorphism. *Genetics* **140**, 783–796 (1995).
30. Kumar, S., Tamura, K. & Nei, M. *MEGA: Molecular Evolutionary Genetics Analysis, version 1.0*. (Pennsylvania State Univ., Univ. Park, 1993).

Acknowledgements

We thank M. Turelli and B. Charlesworth for discussions and advice. D.C. was supported by the Natural Environment Research Council of Great Britain, D.A.F. by a grant from the Leverhulme Trust, and F.M. and I.N. by research contracts from the Centre National de la Recherche Scientifique, the Institut National de la Recherche Agronomique and the Ecole Normale Supérieure de France.

Correspondence and requests for materials should be addressed to D.C. (e-mail: Deborah.Charlesworth@ed.ac.uk).

Representation of a perceptual decision in developing oculomotor commands

Joshua I. Gold & Michael N. Shadlen

Department of Physiology and Biophysics, and Regional Primate Research Center, University of Washington, Seattle, Washington 98195-7290, USA

Behaviour often depends on the ability to make categorical judgements about sensory information acquired over time. Such judgements require a comparison of the evidence favouring the alternatives^{1–4}, but how the brain forms these comparisons is unknown. Here we show that in a visual discrimination task, the accumulating balance of sensory evidence favouring one interpretation over another is evident in the neural circuits that generate the behavioural response. We trained monkeys to make a direction judgement about dynamic random-dot motion⁵ and to indicate their judgement with an eye movement to a visual target. We interrupted motion viewing with electrical microstimulation of the frontal eye field and analysed the resulting, evoked eye movements for evidence of ongoing activity associated with the oculomotor response^{6–10}. Evoked eye movements deviated in the direction of the monkey's judgement. The magnitude of the deviation depended on motion strength and viewing time. The oculomotor signals responsible for these deviations reflected the accumulated motion information that informed the monkey's choices on the discrimination task. Thus, for this task, decision formation and motor preparation appear to share a common level of neural organization.

The difficulty of the direction discrimination task (Fig. 1a, b) depended on the percentage of coherently moving random dots and the amount of time that the monkey was given to view the motion stimulus. At high motion strengths, the monkey can make an accurate direction judgement in a short time. In contrast, when the stimulus is close to the psychophysical threshold, the monkey must base its direction judgement on weak motion signals interpreted over a longer period of time⁵. These judgements are thought to involve the accumulation of motion information represented in the extrastriate visual cortex¹¹. This accumulated motion information ultimately guides selection of the eye-movement response. The fact that the monkey knows beforehand which eye movement it must make to indicate a particular direction judgement suggests that, for this task, its decision can be formed by an ongoing conversion of motion information directly into an appropriate oculomotor command. We tested this idea by using electrically evoked eye movements to assess the influence of motion strength and viewing duration on oculomotor commands that develop

while the monkey is viewing the dots and forming its direction judgement.

We began microstimulation of the frontal eye field (FEF) simultaneously with offset of the motion stimulus and the fixation point, which typically resulted in two, temporally separable eye movements (Fig. 1c–e). The first, electrically evoked movement began 44 ± 10 ms (mean $\pm$ s.d.; 37,611 trials) after the fixation-point offset. This movement probably began before the oculomotor system received the signal to move the eyes voluntarily, as visual latencies exceed 60 ms in the FEF¹² and 40–50 ms in the deep layers of the superior colliculus^{13,14}. The eyes then remained relatively stationary for 95 ± 32 ms, which allowed reliable measurement of the endpoint position of the electrically evoked saccade. The second, voluntary movement began 171 ± 33 ms after fixation-point offset and travelled from the endpoint of the evoked saccade to one of the two choice targets, thereby indicating the monkey's direction judgement.

Eye movements evoked by electrical microstimulation of the FEF during the discrimination task depended on the monkey's subsequent direction judgement. For the experiment depicted in Fig. 2a, the endpoints of all electrically evoked saccades were to the right of the fixation point. However, on trials in which the monkey subsequently chose the upper target, the endpoints of the evoked saccades tended to deviate upwards (filled symbols). Conversely, on trials in which the monkey chose the lower target, the endpoints tended to deviate downwards (open symbols). For all 27 sites tested, saccades

evoked on discrimination trials were similar to those evoked with fixation but tended to deviate in the direction of the subsequently selected target. As shown in Fig. 2b and c, the angle describing this saccade deviation did not differ substantially from the angle defined by the choice targets (mean angular difference $\pm$ s.e.m., $2.8 \pm 3.5^\circ$; paired *t*-test, $P = 0.43$). This result indicates that the trajectory of the electrically evoked saccade was determined by both the site of stimulation and a developing oculomotor command towards the choice target, in a manner consistent with a vector-averaging procedure^{15–17}.

In any one experiment, electrically evoked saccades deviated in the direction of the subsequently selected target by variable amounts (Fig. 2a). The magnitude of this deviation depended on the fraction of coherently moving random dots and the amount of viewing time. Figure 3a shows the endpoints of evoked saccades on trials from a single experiment in which weak motion was presented for a short time, conditions that make it difficult to reach an accurate direction judgement. These evoked saccades varied only weakly with target choice. In contrast, when strong motion was presented for a longer time, evoked saccades from the same site were clearly separated as a function of target choice (Fig. 3b).

The effect of coherence and viewing duration was evident across the population of 37,611 trials from 27 stimulation sites. For each trial, we determined the component of the electrically evoked saccade in the direction of the choice target. On average, this quantity increased in magnitude with increasing motion strength (Fig. 3c; the effect of motion strength was significant for all but the earliest epoch: linear regression, *F* test, $P < 0.01$) and increasing viewing duration (Fig. 3d; the effect of viewing duration was

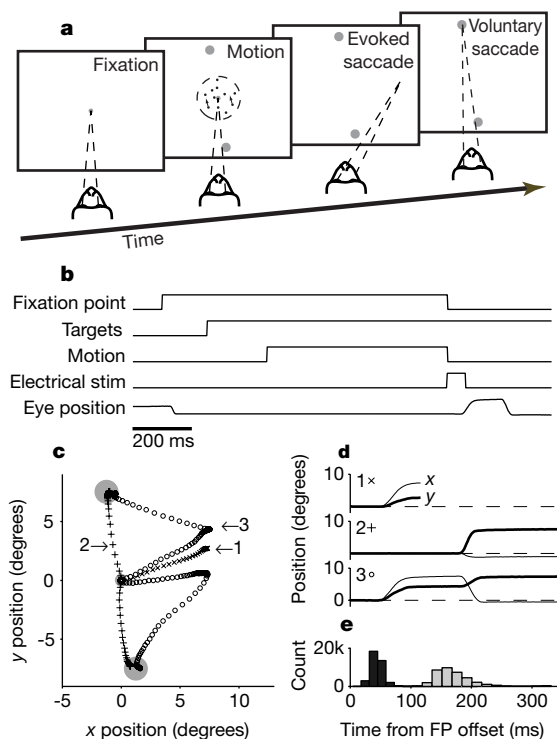


Figure 1 Microstimulation paradigm. **a, b**, After the monkey fixated, two choice targets appeared 8° from the fixation point. The motion stimulus then appeared after 100–600 ms and remained on for 100–1,100 ms, until fixation-point offset. The monkey then indicated its direction judgement with a saccade to one of the targets. On stimulation trials, FEF microstimulation began immediately upon fixation-point offset. **c**, Sample saccade trajectories from single trials (plotted in 2-ms intervals): $\times$, microstimulation following fixation; +, discrimination without microstimulation (two trials); small circles, discrimination with microstimulation (two trials). The two large, grey circles represent the targets. The fixation point is at the origin. **d**, Time courses of the numbered saccades in **c**. Thin and thick lines indicate horizontal (*x*) and vertical (*y*) positions, respectively. **e**, Latencies of first (electrically evoked) and second (voluntary) saccades measured with respect to fixation-point offset for all stimulation trials ($n = 37,611$).

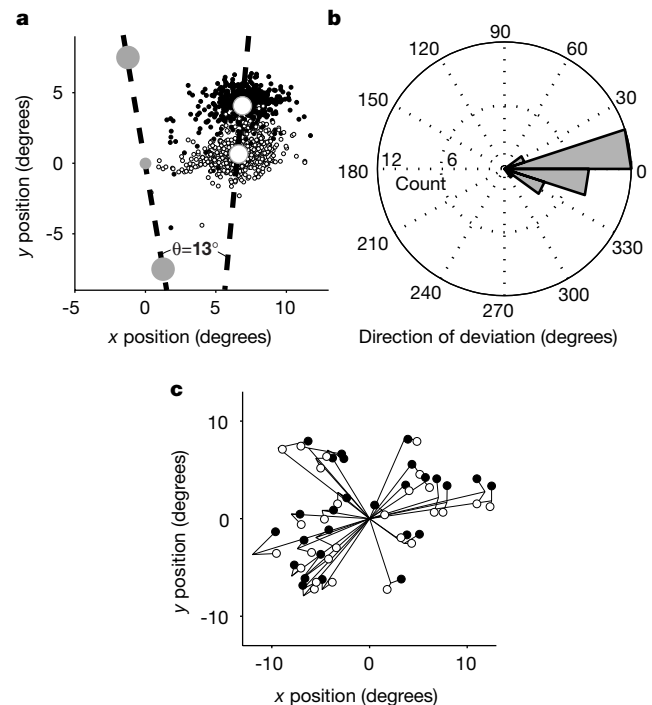


Figure 2 Effect of target choice on eye movements evoked electrically during the discrimination task. **a**, Evoked-saccade endpoints for all correct trials from the experiment shown in Fig. 1. Closed and open symbols represent upward and downward choices, respectively. White circles represent means of the two distributions of endpoints. Dashed lines through the mean endpoints and choice targets describe an angle θ , which compares the direction of saccade deviation to the direction of the motion discrimination. **b**, Distribution of angles (θ) for 27 microstimulation experiments. **c**, Mean endpoint positions of evoked saccades for all correct trials from each site tested. Line segments radiating from the origin represent the mean evoked saccade from fixation trials. Circles show the mean endpoints for evoked saccades from discrimination trials, sorted by the ensuing target choice. Filled circles indicate the more upward choice.

significant at each coherence: polynomial regression, $P < 0.01$). This result implies that dynamic signals present in the oculomotor system during the motion-viewing period reflect the accumulation of motion information over time.

Like the evoked saccades, the monkeys' behavioural performance depended on the strength and viewing duration of the motion stimulus (Fig. 4). We investigated whether the monkeys' performance could be governed by the same process that gives rise to the variation in the evoked saccades. The curves in Fig. 4 represent fits to the behavioural data of a model that predicts the probability of making a correct direction judgement for a given motion strength and viewing duration (equation (2) in Methods). This prediction is based on the difference between two variable quantities, S_c and S_i , which can be thought of as the accumulated responses of motion sensors that encode the correct and incorrect directions of motion, respectively:

$$\begin{aligned} \langle S_c \rangle &= (r_0 + aC^m)T^n \\ \langle S_i \rangle &= r_0 T^n \end{aligned} \quad (1)$$

where $\langle \dots \rangle$ denotes the expectation, C is coherence (0 to 1), T is viewing duration (in s) and r_0 is the response (in spikes s^{-1}) associated with 0% motion coherence. The remaining terms are fitted, positive-valued parameters: a and m parameterize the effects of coherence on the expected signals, and n describes a power law that governs how the signals accumulate in time. S_c and S_i , representing the accumulation of spikes, were treated as random

variables with variances that scaled with their means^{18,19}. For simplicity, only S_c depends on motion strength.

By adjusting the values of the free parameters in equation (1), we found a model that closely predicted the monkeys' ability to make a correct direction discrimination (Fig. 4). For both stimulation and non-stimulation trials, the best fit produced values of m and n that were near unity. These values imply that the monkeys' direction judgements are based on signals that scale in an essentially linear fashion with both coherence and viewing duration. The linear dependence on coherence mimics the responses of neurons in the middle temporal visual area (MT) to random-dot motion²⁰. The linear dependence on time indicates that this type of coherence-dependent activity is temporally integrated to convert spike rate to an accumulated count, similar to the conversion of eye velocity to eye position in the brainstem²¹.

The model of behavioural performance assumes that, on a given trial, the monkeys' decisions are based on the relative magnitudes of the signals defined in equation (1). When S_c is greater, the monkey chooses the actual direction of coherent motion. When S_i is greater, the monkey chooses the incorrect direction. The values of these signals, estimated from the behavioural data, are related to a quantity known as a decision variable, which describes the relative likelihood of one alternative versus another^{1,2,22}. We found that a decision variable in the form of the log of the expected difference between S_c and S_i (equations (3) and (4) in Methods) bears a striking similarity to the magnitude of deviation measured in the evoked saccades (Fig. 5; χ^2 goodness of fit, $Q = 0.047$). For correct choices, both the decision variable and the saccade deviations increased with longer viewing duration and stronger motion. For incorrect choices, both quantities increased with longer viewing duration but decreased with greater motion strength.

The freedom to shift and scale the decision variable derived from performance does not trivialize its close agreement with the magnitude of saccade deviations at all motion strengths. For example, the freedom to adjust each of the solid curves independently did not significantly improve the fit ($P = 0.195$, likelihood ratio test). This result indicates that the coherence dependence of the decision variable is particularly important, even when scaled and shifted. We further assessed the significance of this dependence by fitting the deviation data with a model that was not based on the decision variable. We assumed that oculomotor commands could be initiated with different latencies depending on task difficulty but

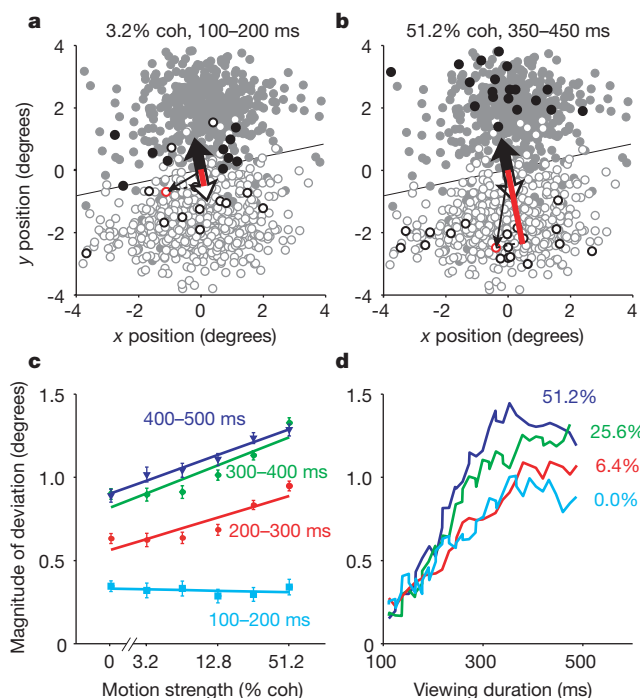


Figure 3 Effect of motion strength and viewing duration on eye movements evoked electrically during the discrimination task. **a, b**, Evoked-saccade endpoints from the experiment in Fig. 2a, corrected for drift during the experiment. The origin represents the average endpoint for both choices (see Methods). The deviation magnitude was quantified as the dot product between the endpoint vector (thin arrows) and the unit vector toward the chosen target (thick arrows). Examples are in red. Black lines indicate zero magnitude. Deviation magnitude was smaller with weak motion and short viewing durations (bold symbols in **a**; mean $\pm$ s.e.m., $0.7 \pm 0.2^\circ$) than with strong motion and longer durations (**b**; $2.4 \pm 0.1^\circ$). **c**, Summary for 27 sites, data separated by viewing duration. Symbols and error bars represent mean $\pm$ s.e.m. **d**, Summary data separated by motion strength. Curves represent 300-point running means. Data in **a–d** are from correct trials.

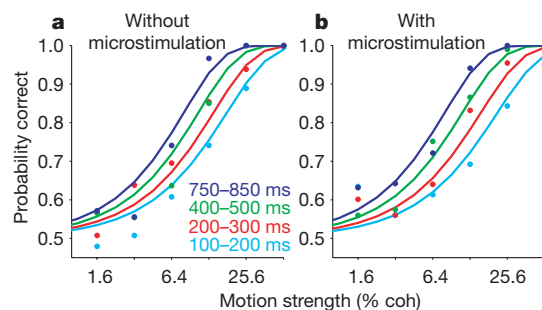


Figure 4 Behavioural performance as a function of motion strength and viewing duration. Viewing duration is binned into four groups, as indicated. Solid curves are maximum-likelihood fits to the behavioural data, derived from equations (1) and (2). **a**, Data from 14,481 non-microstimulation trials. For this fit, $a = 0.37$ spikes s^{-1} per percent coherence (95% confidence interval (CI): 0.29–0.48), $m = 1.04$ (CI: 0.95–1.12) and $n = 0.90$ (CI: 0.75–1.06). **b**, Data from 37,611 microstimulation trials. The monkeys indicated their choices with a voluntary saccade to one of the choice targets after the electrically evoked eye movement. For this fit, $a = 0.37$ spikes s^{-1} per percent coherence (CI: 0.32–0.43), $m = 1.03$ (CI: 0.98–1.13) and $n = 1.09$ (CI: 0.98–1.18). Although the fits in **a** and **b** differed significantly (likelihood-ratio test; $P < 0.0001$), they describe similar effects of motion strength and viewing duration on performance.

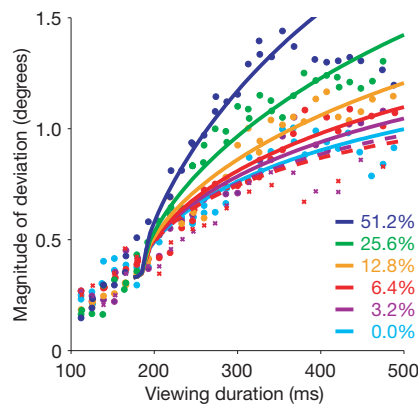


Figure 5 Comparison of the deviation of evoked saccades and a decision variable. Symbols represent the deviations of eye movements evoked electrically during the discrimination task in the direction of the chosen target (see Fig. 3d). These deviations are plotted as a function of viewing duration for different motion strengths, as labelled. Curves represent the logarithm of a decision variable inferred from the behavioural data in Fig. 4b. Filled circles and solid curves represent correct trials; crosses and dashed curves represent error trials.

then developed in a stereotyped manner (equation (5)), as might occur if the decision were computed first. This fit was poor (χ^2 goodness of fit, $Q < 10^{-14}$), unlike the model based on the decision variable (Fig. 5).

Note that for viewing durations of less than 184 ms (δ in equation (4)), motion strength did not affect the evoked saccades even though it did influence the monkeys' choices (Fig. 4). It therefore seems likely that the interpretation of motion information occurred in the time between the evoked and voluntary saccades. The saccade deviations before 184 ms may have reflected a predisposition to choose one of the saccade targets, as seen in physiological recordings from the lateral intraparietal area (LIP) and the superior colliculus^{23,24}. For longer durations, the magnitude of saccade deviation appeared to plateau, perhaps reflecting an end to the linear accumulation of sensory information once a judgement was reached, even if motion viewing continued.

The close relationship between the deviations of evoked saccades and a decision variable inferred from behaviour indicates that developing oculomotor commands may reflect the formation of the monkey's direction judgement. Whether this means that the mechanisms responsible for generating the oculomotor commands are the same as those that generate the judgement is unclear. Decision formation could occur separately from motor preparation, in areas such as the superior colliculus, the prefrontal cortex or area LIP, each of which contains subsets of neurons that appear to be involved in the interpretation of visual motion information^{24–26}. However, some (if not all) of these structures are undoubtedly involved in generating the oculomotor signals reflected in the evoked saccades and thus could take part in a direct conversion of motion information into an eye-movement command. Indeed, the ability to observe intermediate stages of decision formation in oculomotor signals seems implausible if the interpretation of motion information were completed elsewhere, as a fully formed judgement should elicit the same eye movement regardless of the information used to reach that judgement. It remains to be seen whether neural structures associated with the preparation of behavioural responses have a general role in the formation of perceptual judgements. □

Methods

Behavioural task

The motion stimulus was contained within a 5° circular aperture centred on the fixation point (Fig. 1a). On successive video frames, a percentage of randomly chosen dots moved

towards one of two choice targets along the axis of motion discrimination (5° s^{-1}); the remainder of the dots were replotted at random locations. The percentage of coherently moving dots, viewing duration (100 ms plus a random time chosen from an exponential distribution with mean 200 ms) and direction were randomly interleaved. During motion viewing, the monkey maintained fixation within $\pm 0.6^\circ$ and did not make pursuit movements. The monkey received a liquid reward for a correct response (an eye movement to the target in the direction of coherent motion) and on half of the 0% coherence trials.

Electrophysiology

Monkeys were prepared for the microstimulation experiments by surgical implantation of an eye coil, head-holding device and recording cylinder. The FEF was targeted using stereotaxic information and magnetic resonance imaging, as described²⁶. Electrical microstimulation consisted of a train of 0.25-ms-long, biphasic pulses applied at a rate of 350–450 Hz for 60 ms. A site was selected if saccades with consistent trajectories were evoked using $<50 \mu\text{A}$ of current applied in darkness²⁷. The axis of motion discrimination was then rotated to be roughly perpendicular to the trajectory of the evoked saccade, and the microstimulation current was adjusted to evoke saccades reliably during the motion-discrimination task (25–110 μA). In each experiment, 75% of the trials were accompanied by microstimulation.

All training, surgery and experimental procedures were in accordance with the National Institutes of Health Guide for the Care and Use of Laboratory Animals and were approved by the University of Washington Animal Care Committee.

Data analysis

Eye position was monitored using the scleral search-coil technique (sampled at 500 Hz)^{28,29}. On stimulation trials, we included for analysis only trials in which two saccades were observed within 350 ms following microstimulation onset. Saccade endpoints were measured in a 20-ms epoch following the evoked saccade and preceding the monkey's voluntary saccade. Drift in endpoint position over the course of an experiment (lasting 2–4 h) was accounted for by subtracting a 300-point, chronological running mean from the endpoints. The magnitude of saccade deviation was quantified as the dot product between the vector from the fixation point to the saccade endpoint (corrected for drift) and the unit vector from the fixation point towards the target (see Fig. 3a, b). Deviation magnitude was computed only for trials with viewing durations ≤ 500 ms, because measurements from longer trials were sparse.

Fit to behavioural data

According to the model in equation (1), a correct response ensues when $S_c > S_i$. If the signals are independent and normally distributed, then the probability, P , of a correct response is

$$P(C, T) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_0^\infty e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \quad (2)$$

$$\text{where } \mu = \langle S_c \rangle - \langle S_i \rangle, \sigma^2 = (\langle S_c \rangle + \langle S_i \rangle)\phi \text{ and } \phi = \frac{\text{Var}[S_c]}{\langle S_c \rangle} = \frac{\text{Var}[S_i]}{\langle S_i \rangle}$$

The psychometric functions (Fig. 4) were derived by maximizing the (binomial) likelihood of observing the monkeys' choices: that is, $\text{argmax}_{a,m,n} L(\text{choices}|a, m, n)$. The values of ϕ (equation (2)) and r_0 (equation (1)) affect the estimate of a (equation (1)) but cause minimal changes from unity in the values of the exponents, m and n , when fit to the behavioural data. Confidence intervals were estimated by varying the fitted parameters to delimit an acceptance region for the null hypothesis (likelihood ratio test; χ^2 , d.f. = 3, $P > 0.05$). We set $r_0 = 10 \text{ spikes s}^{-1}$, which is consistent with measurements from area MT³⁰, and $\phi = 0.3$, a reasonable value for the pooled responses of weakly correlated neurons^{11,30}. These values yielded $a = 0.37 \text{ spike s}^{-1}$ per percent coherence, which is in close agreement with values reported for area MT²⁰.

Decision variable

The fit of equations (1) and (2) enabled us to derive a decision variable, D , based on a comparison of S_c and S_i :

$$D(C, T) = \begin{cases} \langle S_c - S_i | S_c > S_i \rangle & \text{correct choice} \\ \langle S_i - S_c | S_i > S_c \rangle & \text{incorrect choice} \end{cases} \quad (3)$$

where $\langle x|y \rangle$ indicates the expected value of x under the condition y . D is the expected difference between S_c and S_i , computed separately for correct and incorrect choices. Note that D is always positive.

We fit the measured deviations of electrically evoked saccades with a scaled version of the logarithm of the decision variable, D :

$$f(C, T) = \beta \log[1 + D(C, T - \delta)] + h(T) \quad (4)$$

where β and δ are free parameters representing a scaling factor and a delay, respectively. For $T < \delta$, the fit is dominated by $h(T)$, a time-dependent offset that is modelled as a saturating exponential with a short time constant, $\tau \ll \delta$. The best alignment (least squares) was achieved using $\beta = 0.69$ and $\delta = 184 \text{ ms}$ (Fig. 5). The particular form of the function is probably unimportant beyond the qualitative observation that D , compressed and shifted, matches the family of eye-movement curves obtained by microstimulation. We also fit the

data in Fig. 5 with an alternative model, in which the saccade deviation arises from stereotyped dynamics of oculomotor commands that are simply delayed as a function of motion strength:

$$f(C, T) = \alpha + \beta \left[1 - e^{-\left(\frac{T - \delta_c}{\tau}\right)^n} \right] \quad (5)$$

The free parameters α , β , τ and n describe the temporal dynamics of the saturating exponential function, independent of motion strength. Only the delay term, δ_c , varied with motion strength.

Received 15 October 1999; accepted 26 January 2000.

- Graham, N. V. S. *Visual Pattern Analyzers* (Oxford Univ. Press, Oxford, 1989).
- Carpenter, R. H. & Williams, M. L. Neural computation of log likelihood in control of saccadic eye movements. *Nature* **377**, 59–62 (1995).
- Leon, M. I. & Shadlen, M. N. Exploring the neurophysiology of decisions. *Neuron* **21**, 669–672 (1998).
- Schall, J. D. & Thompson, K. G. Neural selection and control of visually guided eye movements. *Annu. Rev. Neurosci.* **22**, 241–259 (1999).
- Britten, K. H., Shadlen, M. N., Newsome, W. T. & Movshon, J. A. The analysis of visual motion: a comparison of neuronal and psychophysical performance. *J. Neurosci.* **12**, 4745–4765 (1992).
- Schlag-Rey, M., Schlag, J. & Shook, B. Interactions between natural and electrically evoked saccades. I. Differences between sites carrying retinal error and motor error signals in monkey superior colliculus. *Exp. Brain Res.* **76**, 537–547 (1989).
- Kustov, A. A. & Robinson, D. L. Shared neural control of attentional shifts and eye movements. *Nature* **384**, 74–77 (1996).
- Sparks, D. L. & Mays, L. E. Spatial localization of saccade targets. I. Compensation for stimulation-induced perturbations in eye position. *J. Neurophysiol.* **49**, 45–63 (1983).
- Sparks, D. L., Mays, L. E. & Porter, J. D. Eye movements induced by pontine stimulation: interaction with visually triggered saccades. *J. Neurophysiol.* **58**, 300–318 (1987).
- Tehovnik, E. J., Slocum, W. M. & Schiller, P. H. Behavioural conditions affecting saccadic eye movements elicited electrically from the frontal lobes of primates. *Eur. J. Neurosci.* **11**, 2431–2443 (1999).
- Shadlen, M. N., Britten, K. H., Newsome, W. T. & Movshon, J. A. A computational analysis of the relationship between neuronal and behavioral responses to visual motion. *J. Neurosci.* **16**, 1486–1510 (1996).
- Bruce, C. J. & Goldberg, M. E. Primate frontal eye fields. I. Single neurons discharging before saccades. *J. Neurophysiol.* **53**, 603–635 (1985).
- Wurtz, R. H. & Goldberg, M. E. Activity of superior colliculus in behaving monkey. III. Cells discharging before eye movements. *J. Neurophysiol.* **35**, 575–586 (1972).
- Mays, L. E. & Sparks, D. L. Dissociation of visual and saccade-related responses in superior colliculus neurons. *J. Neurophysiol.* **43**, 207–232 (1980).
- Schiller, P. H. & Sandell, J. H. Interactions between visually and electrically elicited saccades before and after superior colliculus and frontal eye field ablations in the rhesus monkey. *Exp. Brain Res.* **49**, 381–392 (1983).
- Lee, C., Rohrer, W. H. & Sparks, D. L. Population coding of saccadic eye movements by neurons in the superior colliculus. *Nature* **332**, 357–360 (1988).
- Sparks, D. L., Lee, C. & Rohrer, W. H. Population coding of the direction, amplitude, and velocity of saccadic eye movements by neurons in the superior colliculus. *Cold Spring Harb. Symp. Quant. Biol.* **55**, 805–811 (1990).
- Geisler, W. S. & Albrecht, D. G. Visual cortex neurons in monkeys and cats: detection, discrimination, and identification. *Vis. Neurosci.* **14**, 897–919 (1997).
- Shadlen, M. N. & Newsome, W. T. The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. *J. Neurosci.* **18**, 3870–3896 (1998).
- Britten, K. H., Shadlen, M. N., Newsome, W. T. & Movshon, J. A. Responses of neurons in macaque MT to stochastic motion signals. *Vis. Neurosci.* **10**, 1157–1169 (1993).
- Robinson, D. A. Integrating with neurons. *Annu. Rev. Neurosci.* **12**, 33–45 (1989).
- Green, D. M. & Swets, J. A. *Signal Detection Theory and Psychophysics* (Wiley, New York, 1966).
- Platt, M. L. & Glimcher, P. W. Neural correlates of decision variables in parietal cortex. *Nature* **400**, 233–238 (1999).
- Horowitz, G. D. & Newsome, W. T. Separate signals for target selection and movement specification in the superior colliculus. *Science* **284**, 1158–1161 (1999).
- Shadlen, M. N. & Newsome, W. T. Motion perception: seeing and deciding. *Proc. Natl Acad. Sci. USA* **93**, 628–633 (1996).
- Kim, J.-N. & Shadlen, M. N. Neural correlates of a decision in the dorsolateral prefrontal cortex of the macaque. *Nature Neurosci.* **2**, 176–185 (1999).
- Bruce, C. J., Goldberg, M. E., Bushnell, M. C. & Stanton, G. B. Primate frontal eye fields. II. Physiological and anatomical correlates of electrically evoked eye movements. *J. Neurophysiol.* **54**, 714–734 (1985).
- Robinson, D. A. A method of measuring eye movement using a scleral search coil in a magnetic field. *IEEE Trans. Biomed. Eng.* **10**, 137–145 (1963).
- Judge, S. J., Richmond, B. J. & Chu, F. C. Implantation of magnetic search coils for measurement of eye position: an improved method. *Vision Res.* **20**, 535–538 (1980).
- Zohary, E., Shadlen, M. N. & Newsome, W. T. Correlated neuronal discharge rate and its implications for psychophysical performance. *Nature* **370**, 140–143 (1994).

Acknowledgements

We thank M. Mihali for animal training and technical support, and E. Freedman, G. Horowitz, B. Jagadeesh, T. Movshon and F. Rieke for helpful comments on the manuscript. This work was supported by the NEI, NCR and the McKnight Foundation.

Correspondence and requests for materials should be addressed to M.N.S. (e-mail: shadlen@u.washington.edu).

A *Drosophila* model of Parkinson's disease

Mel B. Feany* & Welcome W. Bender†

*Department of Pathology, Division of Neuropathology, Brigham and Women's Hospital and Harvard Medical School, and †Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115, USA

Parkinson's disease is a common neurodegenerative syndrome characterized by loss of dopaminergic neurons in the substantia nigra, formation of filamentous intraneuronal inclusions (Lewy bodies) and an extrapyramidal movement disorder. Mutations in the α -synuclein gene are linked to familial Parkinson's disease^{1,2} and α -synuclein accumulates in Lewy bodies and Lewy neurites^{3–5}. Here we express normal and mutant forms of α -synuclein in *Drosophila* and produce adult-onset loss of dopaminergic neurons, filamentous intraneuronal inclusions containing α -synuclein and locomotor dysfunction. Our *Drosophila* model thus recapitulates the essential features of the human disorder, and makes possible a powerful genetic approach to Parkinson's disease.

It is unclear how mutations in α -synuclein, an abundant neuronal protein of unknown function, produce neurodegeneration in familial cases of Parkinson's disease. However, the dominant inheritance pattern and production of insoluble protein aggregates indicate a toxic dominant mechanism, perhaps relating to abnormal protein accumulation. Expression of human α -synuclein in *Drosophila* might therefore model Parkinson's disease. We have produced transgenic fly lines that produce normal human α -synuclein and separate lines with each of the two mutant proteins linked to familial Parkinson's disease, A30P and A53T α -synuclein.

We use a bipartite expression system that relies on transcriptional activation by the yeast protein GAL4 (ref. 6) to express normal and mutant α -synuclein in flies. Normal and mutant human α -synuclein complementary DNA constructs are placed downstream of multiple binding sites for GAL4. Transgenic animals carrying the GAL4-responsive constructs are then crossed to a number of well characterized lines that express the yeast activator in a variety of tissue- and cell-type-specific patterns (the 'drivers'). Development of neuronal and non-neuronal tissues proceeds normally in the presence of human α -synuclein, as indicated by appropriate external morphology, histological appearance (Fig. 1a, b), viability and behaviour of newly eclosed flies (Table 1). To confirm appropriate activity of the drivers, we expressed an unrelated toxic protein, mutant ataxin-1, in the same developmental and tissue-specific patterns. In contrast to α -synuclein, mutant ataxin-1 produces marked defects during development (Table 1). Other laboratories have also induced phenotypes with these driver lines⁷.

The nervous system is appropriately formed in α -synuclein transgenic flies, and the aging brain shows no widespread degenerative changes (Fig. 1a, b). Overall brain volume is likewise preserved in many patients with Parkinson's disease, but marked degeneration occurs in specific groups of neurons. Dopaminergic neurons are preferentially lost in Parkinson's disease; we therefore examined dopaminergic neurons in α -synuclein transgenic flies. We used multiple independent markers to identify dopaminergic cells.

First, we stained serial tissue sections through the entire brain of the adult fly with an antibody against tyrosine hydroxylase, which specifically identifies dopaminergic neurons. The position and arrangement of dopaminergic neurons in the developing and adult *Drosophila* nervous system has been documented in detail^{8–10}. Tyrosine hydroxylase immunostaining shows a prominent cluster of dopaminergic neurons (the dorsomedial group) that is

always represented in well oriented sections by four or five robustly staining cells (Fig. 1c). These cells are present in one-day-old control flies and are not lost with age, being readily identified in 60-day-old flies. The lifespan of control flies and α -synuclein transgenic animals is about 60 days under our 25 °C culture conditions.

In contrast, *Drosophila* that express α -synuclein in a pan-neural distribution (*elav*-GAL4 driver) show a marked, age-dependent loss of dorsomedial dopaminergic neurons (Fig. 1d). In young adult flies expressing wild-type and mutant forms of α -synuclein, the dorsomedial cluster consists of the normal complement of four or five neurons. In 30–60-day-old α -synuclein transgenic flies, however, the cluster is either absent or consists of a single tyrosine-hydroxylase-positive cell (Fig. 1d). These dopaminergic neurons disappear in flies expressing wild-type, A30P or A53T α -synuclein.

To confirm that dopaminergic neurons degenerate in α -synuclein transgenic animals, we used a driver line containing the promoter for the DOPA decarboxylase gene (*Ddc*-GAL4), another marker of dopaminergic neurons. The *Ddc* line drives reporter-gene expression specifically in dopaminergic cells, allowing us to identify them. Transgenic flies expressing A30P α -synuclein under the control of the *Ddc* driver line show robust immunostaining for α -synuclein at one day of age. However, after 30 days, there is no α -synuclein immunostaining associated with cell bodies, consistent with degeneration of dopaminergic neurons. To ensure that the *Ddc* driver line remains active at 30 days, the marker protein β -galactosidase was expressed instead of α -synuclein. β -galactosidase immunoreactivity is present at 30 days when the *Ddc* line is used to drive β -galactosidase expression. We also confirmed that dorsomedial cluster neurons are depleted at 30 days by tyrosine hydroxylase immunostaining brains from flies expressing wild-type, A30P or A53T α -synuclein under the control of the *Ddc*-GAL4 driver.

As a final test for degeneration of dopaminergic neurons, both β -galactosidase and A30P α -synuclein were expressed at the same

time, in the same dopaminergic neurons, using the *Ddc* driver line. In one-day-old transgenic flies, dopaminergic neurons are readily identified by β -galactosidase immunostaining. In contrast, at 10 days of age, β -galactosidase expression is undetectable (Fig. 1f, compare with 10-day-old controls in Fig. 1e).

Not all dopaminergic neurons degenerate in α -synuclein transgenic flies. Substantial numbers of tyrosine-hydroxylase-positive cells remain in aged flies that express α -synuclein in a pan-neural distribution (*elav*-GAL4 driver) or specifically in dopaminergic neurons (*Ddc*-GAL4 driver). Preserved neurons may reflect variation in the amount of GAL4 activator protein available to drive α -synuclein expression in particular cells. Alternatively, certain subsets of *Drosophila* dopaminergic neurons may be particularly sensitive to α -synuclein toxicity. In patients with Parkinson's disease, preferential degeneration of specific dopaminergic neurons occurs even within the same neuronal nucleus¹¹.

Neurodegeneration induced by α -synuclein shows specificity for dopaminergic neurons. Pan-neural expression of α -synuclein in the brain produces no demonstrable loss of volume in the outer cellular cortex or in the central neuropil area (Fig. 1b). The excess vacuolization characteristic of other, more generalized, neurodegenerative mutations in *Drosophila* is not present^{12,13}. In addition, no excess of degenerating neurons is detected by toluidine blue staining or ultrastructural examination. Thus, most neurons are preserved in flies expressing α -synuclein. Of course, such anatomical investigations do not exclude degeneration of a minor neuronal subpopulation.

To address the possibility that subsets of non-dopaminergic neurons are vulnerable to α -synuclein toxicity, we examined a second major amine transmitter in *Drosophila*, serotonin. Serotonergic neurons can degenerate in patients with Parkinson's disease¹⁴. Whole-mount preparations of adult brains stained with an antibody against serotonin revealed no differences between 30-day-old flies expressing A30P α -synuclein in a pan-neural pattern and controls.

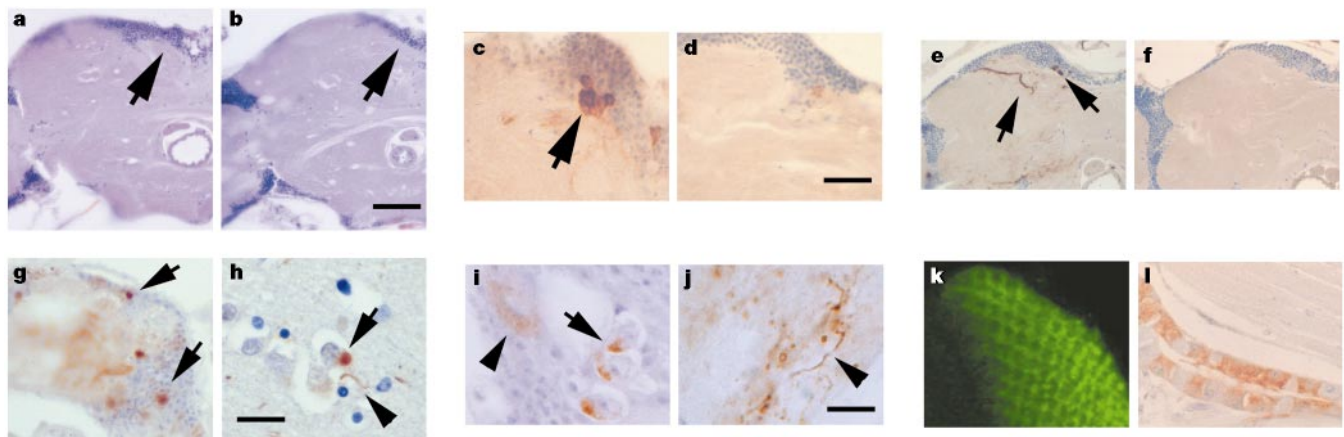


Figure 1 Histological and immunocytochemical analysis of α -synuclein transgenic flies. **a,b**, Frontal sections of 60-day-old control fly (**a**, *elav*-GAL4/+), and 60-day-old A30P α -synuclein transgenic fly (**b**, *UAS*-A30P α -synuclein/*elav*-GAL4) stained with haematoxylin and eosin. Overall brain volume, including the outer cellular cortex layer containing neuronal and glial cell bodies (arrows) and central neuropil areas, and overall architecture are preserved. **c**, 30-day-old control fly (*elav*-GAL4/+) shows immunostaining for tyrosine hydroxylase in 4–5 cells in the dorsomedial cluster. **d**, 30-day-old α -synuclein-expressing fly (*elav*+/+; *UAS*-wild-type α -synuclein/+) shows no cell-body-associated immunostaining in the same area. **e**, 10-day-old control fly expressing α -galactosidase in both cell cortex and neuronal processes (arrows) of dopaminergic neurons (*UAS*-*lacZ*/*Ddc*-GAL4). **f**, 10-day-old fly carrying in addition an α -synuclein transgene (*UAS*-A30P α -synuclein/+/+; *UAS*-*lacZ*/*Ddc*-GAL4) shows no α -galactosidase expression in the outer cellular cortex or central neuropil. **g**, Immunostaining of α -synuclein inclusions in the brain from a 30-day-old transgenic fly (*UAS*-A30P

α -synuclein/*elav*-GAL4) in the area of the subesophageal ganglia with an antibody against α -synuclein. **h**, Human cortical Lewy body (arrow) from the cingulate cortex of a patient with diffuse Lewy body disease, stained with an antibody against ubiquitin (same scale as **g**). **i**, Immunostaining of three α -synuclein inclusions in the brain of a young adult fly (1 day post-eclosion; *Ddc*-GAL4/*UAS*-wild-type α -synuclein) showing irregularity of selected inclusions (arrow) and diffuse immunoreactivity in a larger neuron (arrowhead). **j**, Neuritic pathology consisting of α -synuclein immunoreactive thread- (arrowhead) and grain-like structures. 60-day-old fly (*UAS*-A30P α -synuclein/*elav*-GAL4). Compare with abnormal neurite in the cingulate cortex of a Lewy body disease patient (**h**, arrowhead). **k**, Immunofluorescence staining of eye imaginal disc from wandering third instar larva (*UAS*-A53T α -synuclein/+/+; *gmr*-GAL4/+) with antibody against α -synuclein showing no inclusions. **l**, Diffuse cytoplasmic α -synuclein immunoreactivity in the adult gut from a 30-day-old fly (*UAS*-A53T α -synuclein/+/+; *e29c*-GAL4/+/+). Scale bars: **a,b,e,f**, 30 μ m; **c,d,g,h**, 10 μ m; **i,j**, 5 μ m.

The anatomical arrangement of serotonergic cells has been described in detail in *Drosophila*¹⁵. We can identify all the major serotonergic cell groups in experimental animals and controls. However, we cannot exclude an effect of α -synuclein expression on a small subgroup of serotonergic cells.

Expression of human α -synuclein in flies thus replicates three key features of the pathology of Parkinson's disease: adult onset, involvement restricted to the nervous system and anatomical specificity within the nervous system.

The most specific and diagnostic feature of Parkinson's disease is an α -synuclein-rich cytoplasmic filamentous aggregate called the Lewy body. Lewy bodies in the cerebral cortex are best visualised by immunostaining^{16,17}. When we immunostain brains from aged flies expressing normal and mutant α -synuclein in a pan-neural pattern with antibodies against human α -synuclein, we observe a distinct punctate pattern of staining suggestive of aggregate formation (Fig. 1g). The fly α -synuclein inclusions strongly resemble cortical Lewy bodies from patients with diffuse Lewy body disease, a disorder closely related to Parkinson's disease (Fig. 1h). Most neuronal cytoplasmic inclusions are single, round and regular, as in Fig. 1g. Occasional inclusions are more irregular (Fig. 1i). Neuritic α -synuclein pathology is also present in transgenic fly brains, and consists of both thread- and grain-like inclusions (Fig. 1j). The light microscopic morphology of inclusions produced by wild-type, A30P or A53T α -synuclein transgenic flies is indistinguishable.

Electron microscopy reveals cytoplasmic inclusions in brains from flies expressing α -synuclein (pan-neural *elav*-GAL4 driver) that have a relatively homogenous core and are edged by radiating filaments projecting into a surrounding halo (Fig. 2a). The filaments are 7–10 nm in diameter and are sometimes associated with cellular organelles (Fig. 2b, arrow). Granular material is admixed with loosely packed filaments in less compact inclusions (Fig. 2b, arrowhead), as in cortical Lewy bodies^{16–18}. The overall morphology of the inclusions, the filamentous and granular nature of the aggregates and the size and disposition of the component filaments are all reminiscent of human Lewy bodies¹⁸. We never saw similar inclusions in aged control flies. Immunoelectron microscopy reveals Hrp labelling concentrated over the inclusions (Fig. 2c). Faint radiating filaments are visible peripherally in the unstained preparation (Fig. 2c, arrow). No immunoreactivity was present in identically treated specimens from control flies of the same age.

The time of appearance of α -synuclein inclusions in the fly brain depends on the level of α -synuclein expression. The inclusions appear at 20–30 days of age when α -synuclein expression is pan-neural. In younger brains, the immunostaining pattern is diffuse and cytoplasmic (Fig. 1i, arrowhead). Immunoreactivity is also

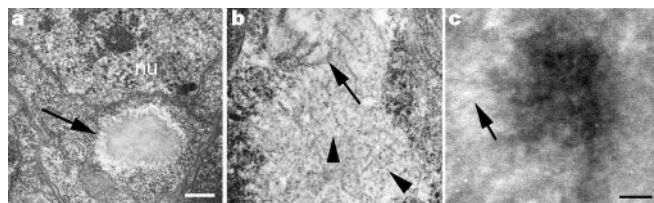


Figure 2 Electron microscopy of α -synuclein inclusions in flies. **a**, Neuronal cytoplasmic inclusion with relatively homogenous core and peripheral radiating filaments (arrow) from a 25-day-old α -synuclein transgenic fly (*UAS-A30P α -synuclein/elav-GAL4*). nu, nucleus. **b**, Higher magnification of a more diffuse inclusion from a fly of the same age and genotype showing loosely arranged filaments, sometimes associated with cellular organelles (arrow), intermixed with granular material (arrowhead). **c**, Immunoelectron microscopy showing central electron-dense Hrp immunolabelling in an inclusion from a 60-day-old fly of the same genotype. Faint radiating filaments (arrow) are visible peripherally in this unstained specimen. Scale bars: **a**, 0.5 μ m; **b**, **c**, 0.1 μ m.

diffuse and cytoplasmic in developing tissues (Fig. 1k, eye imaginal disc) and non-neuronal tissues (Fig. 1l, adult gut) when the GAL4 drivers induce expression in these tissues. Inclusion formation thus parallels α -synuclein toxicity in both its restriction to the nervous system and its timing. Inclusions are present in 2–10% of neurons throughout the central body complex at 30 days of age (pan-neural *elav*-GAL4 driver), and are not restricted to dopaminergic neurons. Lewy bodies in classic Parkinson's disease and in diffuse Lewy body disease are also present in a variety of nuclei and neuronal subtypes, not all of which show obvious degeneration.

Given the neuronal degeneration and widespread inclusion formation present in α -synuclein transgenic flies, we searched for behavioural manifestations of nervous system dysfunction. Locomotor behaviour is grossly preserved in young flies. However, as transgenic flies with pan-neural expression of α -synuclein age, they develop locomotor dysfunction (Fig. 3). Normal *Drosophila* display a strong negative geotactic response. When tapped to the bottom of a vial they rapidly climb to the top of the vial, and most flies remain there. As they get older, normal flies no longer climb to the top of the vial, but instead make short abortive climbs and fall back to the bottom of the vial. The loss of the climbing response has been used to monitor aging-related changes in *Drosophila*^{19,20}. Flies transgenic for α -synuclein initially climb as well as control flies. However, over time they decline in performance more rapidly than controls (Fig. 3). The progressive, accelerated decline in climbing ability in α -synuclein transgenic flies demonstrates a functional deficit produced by α -synuclein expression in the nervous system. The time course of locomotor dysfunction parallels the degeneration of dopaminergic neurons and the appearance of α -synuclein inclusions.

Drosophila expressing A30P α -synuclein lose their climbing ability earlier than flies expressing wild-type or A53T α -synuclein. This small but statistically significant effect may reflect either enhanced biological toxicity of the A30P mutant protein or small variations in the amount of α -synuclein produced in the transgenic lines. We examined α -synuclein expression in multiple independent transgenic lines for each α -synuclein variant by western blotting, by immunofluorescence on eye imaginal discs and by immunohistochemical staining on sections of adult *Drosophila*, and we have

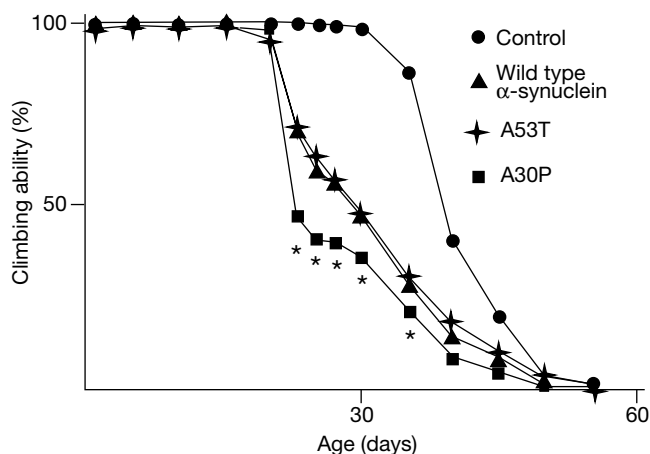


Figure 3 Premature loss of climbing ability in α -synuclein transgenic flies. Flies expressing wild-type, A30P and A53T α -synuclein are significantly different from control flies from days 23 to 45 ($P < 0.01$, one-way analysis of variance with supplementary Newman–Keuls test). Asterisk, climbing scores for A30P transgenic flies that are significantly different from the wild-type and A53T transgenic scores. The s.e.m.s of 20 trials are within the symbols. Control genotype: *elav-GAL4/+*. Experimental genotypes: (1) *elav/+; UAS-wild-type α -synuclein/+*; (2) *UAS-A30P α -synuclein/elav-GAL4*; (3) *UAS-A53T α -synuclein/elav-GAL4*.

Table 1 Effects of expressing normal and mutant α -synuclein in *Drosophila*

Transgene	Tissue and stage transgene expressed				
	Nervous system	Eye	Presumptive mesoderm	Embryo	Imaginal disc, adult tissues
Wild-type α -synuclein	Neurodegeneration	Retinal degeneration	Viable	Viable	Viable
A30P α -synuclein	Neurodegeneration	Retinal degeneration	Viable	Viable	Viable
A53T α -synuclein	Neurodegeneration	Retinal degeneration	Viable	Viable	Viable
SCA1-82	Lethal	Rough eye	Lethal	Lethal	Crumpled wing lethality

SCA1-82, spinocerebellar ataxia type 1 with an expanded repeat length encoding 82 glutamines. GAL4 enhancer lines with the following patterns were used: *elav*-GAL4, all neurons of the developing and mature central and peripheral nervous system; *gmr*-GAL4, developing and mature retina, including photoreceptors and pigment cells; *24B*-GAL4, presumptive mesoderm and muscle; *32B*-GAL4, widespread embryonic, eye and wing imaginal disc; *dpp*-GAL4, imaginal disc anterior-posterior border; *e29c*-GAL4, widespread imaginal disc and adult imaginal disc derivatives. *UAS*-SCA1-82/+; *dpp*-GAL4/+ transheterozygotes showed crumpled wings; animals of the genotype *UAS*-SCA1-82/+; *e29c*-GAL4/+ died as larvae.

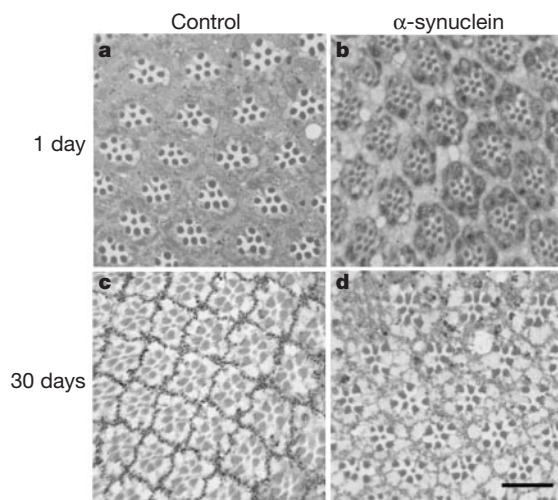


Figure 4 Retinal degeneration in α -synuclein transgenic flies. **a**, One-day-old control fly. **b**, Normal retina in one-day-old α -synuclein transgenic fly. **c**, Well preserved retina in 30-day-old control fly. **d**, 30-day-old α -synuclein-expressing fly showing retinal degeneration with vacuolization and architectural distortion. Scale bar, 15 μ m. Genotypes include control: *gmr*-GAL4/+; transgenic: *UAS*-wild-type α -synuclein/*gmr*-GAL4 flies.

compared transgenic lines with similar levels of α -synuclein as determined by all three assays. However, small differences in α -synuclein protein levels may have escaped our detection.

Degenerative changes are not restricted to the brain. Retinal degeneration occurs when α -synuclein is expressed specifically in the eye (*gmr*-GAL4 driver). Expression of wild-type or mutant α -synuclein during development of the eye produces no effect. However, continued expression of α -synuclein in the adult eye produces retinal degeneration that is detectable by ten days, and marked at 30 days in transgenic flies expressing wild-type (Fig. 4), A30P or A53T α -synuclein. Retinal degeneration can be readily monitored under the dissecting microscope in live flies by examining an optical effect termed the retinal pseudopupil²¹. The pseudopupil is quite sensitive to disruptions in the normal architecture of the retina, and becomes abnormal as the retinas of the α -synuclein transgenic flies degenerate. The presence of an easily assayed, nervous-system-specific, degenerative phenotype will facilitate generation of second site modifiers and other pharmacological and transgenic manipulations designed to modify neurodegeneration.

Retinal degeneration in α -synuclein transgenic flies shows that in flies, as in people, α -synuclein-related degenerative changes show relative rather than absolute specificity for dopaminergic neurons. In fact, the *Drosophila* retina may be resistant to α -synuclein toxicity because more transgenic α -synuclein is produced in the eye (*gmr*-GAL4) than the brain (*elav*-GAL4).

The modest difference between the toxicity of wild-type and A30P mutant α -synuclein in the locomotor assay, and the similar effects of all three α -synuclein variants on dopaminergic neuronal

loss and retinal degeneration, are consistent with the fact that most patients with Parkinson's disease have the wild-type form of the protein, which does become incorporated into Lewy bodies. We cannot show that A53T mutant α -synuclein is more toxic than the wild-type protein in any of our assays. A53T mutant α -synuclein shows accelerated aggregation *in vitro*^{22,23} and enhanced toxicity in some assays^{24,25}. The mouse, however, carries the A53T allele as its normal sequence²⁶. A53T mutant α -synuclein may be better tolerated in animals than is suggested by *in vitro* and cellular assays.

We have created a model of Parkinson's disease by expressing human α -synuclein in *Drosophila*. The parallel time course of dopaminergic cell degeneration, inclusion formation and locomotor dysfunction indicates that the three abnormalities may be causally related. We will now be able to delineate underlying pathogenetic mechanisms and identify novel proteins mediating α -synuclein toxicity in a genetically tractable organism. Our transgenic fruit flies provide a genetic complement to α -synuclein transgenic mice²⁷. □

Methods

Transgenic *Drosophila*

We cloned normal and mutant α -synuclein and SCA1 cDNAs into the GAL4-responsive pUAST expression vector, and created transgenic strains by embryo injection. At least four independent transgenic lines were derived for wild-type α -synuclein and each α -synuclein mutant, and all were tested for protein expression. Western blots with two monoclonal antibodies (LB509 and clone 42, see below) revealed an appropriately sized band (*M*_r 19K) in heads of transheterozygous flies of the following six genotypes: (1) *elav*-GAL4/+; *UAS*-wild-type α -synuclein/+; (2) *UAS*-A30P α -synuclein/*elav*-GAL4; (3) *UAS*-A53T α -synuclein/*elav*-GAL4; (4) *UAS*-wild-type α -synuclein/*gmr*-GAL4; (5) *UAS*-A30P α -synuclein/+; *gmr*-GAL4/+; (6) *UAS*-A53T α -synuclein/+; *gmr*-GAL4/+ . Wild-type and mutant lines expressed 30% or less α -synuclein per mg total brain protein compared with rat and human controls. For wild-type α -synuclein and each mutant, we analysed 2–5 independent transgenic lines with varying α -synuclein expression in the assays described.

The GAL4 driver lines used include: *24B*-GAL4 (ref. 6), *32B*-GAL4 (ref. 6), *dpp*-GAL4, *Ddc*-GAL4 (ref. 28), *e29c*-GAL4, *elav*-GAL4 (ref. 29) and *gmr*-GAL4 (ref. 30). Flies expressing GAL4 from the *Ddc*, *elav* or *gmr* promoter, but without a *UAS*- α -synuclein target, were used as controls, and had the following genotypes: *Ddc*-GAL4/+; *elav*-GAL4/+ and *gmr*-GAL4/+ . The wild-type chromosome from the control heterozygotes was derived from the *w*- background strain used for embryo injections. The GAL4-responsive lacZ reporter construct was *UAS*-lacZ⁸⁹⁴⁻¹⁻² (ref. 6).

We used a *Ddc*-GAL4 line to express α -synuclein and β -galactosidase in dopaminergic cells²⁸. At the time points reported, in flies of the genotype *Ddc*-GAL4/*UAS*-lacZ⁸⁹⁴⁻¹⁻², all cells clearly positive for β -galactosidase also expressed tyrosine hydroxylase by double-label immunohistochemistry. Only a subset of tyrosine-hydroxylase-positive cells expressed β -galactosidase at our level of detection. We thus restrict our conclusions to a subset of dopaminergic cells in experiments with the *Ddc*-GAL4 line.

Sectioning, immunostaining and electron microscopy

Adult flies were fixed in formalin at 1, 10, 30 and 60 days, and embedded in paraffin. Whole-mount preparations fixed in formalin were also used. Immunostaining on paraffin sections was performed using an avidin-biotin-peroxidase complex (ABC) method. Antibodies, sources and dilutions include: anti-tyrosine hydroxylase, Chemicon, 1:500; LB509, Zymed Laboratories, 1:500; clone 42, Transduction Laboratories, 1:200; anti-serotonin, Sigma, 1:500; anti- β -galactosidase, Promega, 1:500; anti-ubiquitin polyclonal serum, Chemicon, 1:1,000. No endogenous immunoreactivity was revealed in tissue sections from nontransgenic control *Drosophila* stained with the anti- α -synuclein antibodies LB509 and clone 42.

To assess brain morphology, serial 4- μ m sections including the entire brain were prepared from formalin-fixed, paraffin-embedded heads from flies of the following three

genotypes: (1) *elav-GAL4/+; UAS-wild-type α -synuclein/+*; (2) *UAS-A30P α -synuclein/elav-GAL4*; and (3) *UAS-A53T α -synuclein/elav-GAL4*. Sections were stained with haematoxylin and eosin. Time points monitored were 1, 10, 30 and 60 days. In addition, serial 1- μ m sections of glutaraldehyde-fixed heads from flies of the same genotypes prepared at 1, 30 and 60 days were stained with toluidine blue to highlight degenerating cells. No evidence of excess neurodegeneration was detected using either technique.

To evaluate dopaminergic cells of the dorsomedial cluster by tyrosine hydroxylase immunostaining, serial 4- μ m sections were cut to include the entire brain. Immunopositive cells at the level of the giant interneuron commissure, posterior to the fan-shaped body, were counted in well oriented frontal sections at 1, 10, 30 and 60 days. At 1 day all control and experimental sections contained four or five cells in the delineated region. At 30 and 60 days all controls showed four or five cells. At 30 and 60 days all α -synuclein-expressing animals (α -synuclein, *elav-GAL4* and α -synuclein, *Ddc-GAL4* transheterozygotes) showed 0 or 1 tyrosine-hydroxylase-positive cell in the defined region. Tyrosine-hydroxylase-positive cells outside the dorsomedial cluster were present, and served as internal controls for the immunostaining procedure. At least four, and usually between six and ten brains were examined for wild-type α -synuclein and each mutant α -synuclein. Controls included young and aged flies of the genotypes *elav-GAL4/+* and *Ddc-GAL4/+*. We evaluated expression of α -synuclein and β -galactosidase on similar serial section preparations. Quantification was simplified in these experiments because no clear cell-body-associated α -synuclein or β -galactosidase immunoreactivity was observed in the aged α -synuclein transgenic flies at the times reported.

For histological examination of retinas, heads were fixed in glutaraldehyde and embedded in epon. Tangential retinal sections were prepared at a thickness of 1 μ m and stained with toluidine blue (Fig 4).

Standard electron microscopy was performed on brains from 25-day-old experimental (*UAS-A30P α -synuclein/elav-GAL4*) and control (*elav-GAL4/+*) flies. For immunoelectron microscopy, pre-embedding immunohistochemistry with an Hrp-conjugated secondary antibody was performed on 60-day adult brains from experimental (*UAS-A30P α -synuclein/elav-GAL4*) and control (*elav-GAL4/+*) flies fixed in 4% paraformaldehyde with 0.5% glutaraldehyde. Tissue was post-fixed in osmium and embedded in epon. Unstained ultrathin sections and ultrathin sections stained with uranyl acetate and lead citrate were examined.

Climbing assay

The climbing assay was performed as described^{19,20}. Forty flies were placed in a plastic vial, and gently tapped to the bottom of the vial. The number of flies at the top of the vial was counted after 18 s of climbing. Twenty trials were performed for each time point. The data shown represent results from a cohort of flies tested serially over 55 days. The experiment was repeated three times, with independently derived transgenic lines. Similar results were obtained from each experiment. The experiment was carried out under red light (Kodak Safelight Filter 1A). Control flies were of the genotype *elav-GAL4/+*. Experimental animals were of the following genotypes: (1) *elav-GAL4/+; UAS-wild-type α -synuclein/+*; (2) *UAS-A30P α -synuclein/elav-GAL4*; and (3) *UAS-A53T α -synuclein/elav-GAL4*.

Received 14 February; accepted 2 March 2000.

1. Polymeropoulos, M. H. *et al.* Mutation in the α -synuclein gene identified in families with Parkinson's disease. *Science* **276**, 2045–2047 (1997).
2. Krüger, R. *et al.* Ala30Pro mutation in the gene encoding α -synuclein in Parkinson's disease. *Nature Genet.* **18**, 106–108 (1998).
3. Spillantini, M. G., Schmidt, M. L., Lee, V. M. -Y. & Trojanowski, J. Q. α -Synuclein in Lewy bodies. *Nature* **388**, 839–840 (1997).
4. Spillantini, M. G., Crowther, R. A., Jakes, R., Hasegawa, M. & Goedert, M. α -Synuclein in filamentous inclusions of Lewy bodies from Parkinson's disease and dementia with Lewy bodies. *Proc. Natl Acad. Sci. USA* **95**, 6469–6473 (1998).
5. Baba, M. *et al.* Aggregation of α -synuclein in Lewy bodies of sporadic Parkinson's disease and dementia with Lewy bodies. *Am. J. Pathol.* **152**, 879–884 (1998).
6. Brand, A. H. & Perrimon, N. Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* **118**, 401–415 (1993).
7. Warrick, J. M. *et al.* Expanded polyglutamine protein forms nuclear inclusions and causes neural degeneration in *Drosophila*. *Cell* **93**, 939–949 (1998).
8. Budnik, V., Martin-Morris, L. & White, K. Perturbed pattern of catecholamine-containing neurons in mutant *Drosophila* deficient in the enzyme dopa decarboxylase. *J. Neurosci.* **6**, 3682–3691 (1986).
9. Budnik, V. & White, L. Catecholamine-containing neurons in *Drosophila melanogaster*: Distribution and development. *J. Comp. Neurol.* **268**, 400–413 (1988).
10. Lundell, M. J. & Hirsh, J. Temporal and spatial development of serotonin and dopamine neurons in the *Drosophila* CNS. *Dev. Biol.* **165**, 385–396 (1994).
11. Gibbs, W. R. G. & Lees, A. J. Anatomy, pigmentation, ventral and dorsal subpopulations of the substantia nigra, and differential cell death in Parkinson's disease. *J. Neurol. Neurosurg. Psychiatry* **54**, 388–396 (1991).
12. Coombe, P. E. & Heisenberg, M. The structural brain mutant *Vacuolar medulla* of *Drosophila melanogaster* with specific behavioral defects and cell degeneration in the adult. *J. Neurogenet.* **3**, 135–158 (1986).
13. Buchanan, R. L. & Benzer, S. Defective glia in the *Drosophila* brain degeneration mutant *drop-dead*. *Neuron* **10**, 839–850 (1993).
14. Jellinger, K. A. Pathology of Parkinson's disease. Changes other than the nigrostriatal pathway. *Mol. Chem. Neuropathol.* **14**, 153–197 (1991).
15. Valls, A. M. & White, K. Serotonin-containing neurons in *Drosophila melanogaster*: Development and distribution. *J. Comp. Neurol.* **268**, 414–428 (1988).
16. Kuzuhara, S., Mori, H., Izumiya, N., Yoshimura, M. & Ihara, Y. Lewy bodies are ubiquitinated: a light and electron microscopic immunocytochemical study. *Acta Neuropathol.* **75**, 345–353 (1988).

17. Dickson, D. W. *et al.* Diffuse Lewy body disease: light and electron microscopic immunocytochemistry of senile plaques. *Acta Neuropathol.* **78**, 572–584 (1987).
18. Pollanen, M. S., Dickson, D. W. & Bergeron, C. Pathology and biology of the Lewy body. *J. Neuropathol. Exp. Neurol.* **52**, 183–191 (1993).
19. Ganetzky, B. & Flanagan, J. R. On the relationship between senescence and age-related changes in two wild-type strains of *Drosophila melanogaster*. *Exp. Gerontol.* **13**, 189–196 (1978).
20. Le Bourg, E. & Lints, F. A. Hypergravity and aging in *Drosophila melanogaster*. 4. Climbing activity. *Gerontology* **38**, 59–64 (1992).
21. Franceschini, N. In *Information Processing in the Visual System of Drosophila* (ed. Wehner, R.) 75–82 (Springer, Berlin, 1972).
22. Conway, K. A., Harper, J. D. & Lansbury, P. T. Accelerated *in vitro* fibril formation by a mutant α -synuclein linked to early-onset Parkinson disease. *Nature Med.* **4**, 1318–1320 (1998).
23. Narhi, L. *et al.* Both familial Parkinson's disease mutations accelerate α -synuclein aggregation. *J. Biol. Chem.* **274**, 9843–9846 (1999).
24. El-Agnaf, O. M. A. *et al.* Aggregates from mutant and wild-type α -synuclein proteins and NAC peptide induce apoptotic cell death in human neuroblastoma cells by formation of β -sheet and amyloid-like filaments. *FEBS Lett.* **440**, 71–75 (1998).
25. Ostreova, N. *et al.* α -Synuclein shares physical and functional homology with 14-3-3 proteins. *J. Neurosci.* **19**, 5782–5791 (1999).
26. Hong, L. *et al.* The cDNA cloning and ontogeny of mouse α -synuclein. *NeuroReport* **9**, 1239–1243 (1998).
27. Masliah, E. *et al.* Dopaminergic loss and inclusion body formation in α -synuclein mice: implications for neurodegenerative disorders. *Science* **287**, 1265–1269 (2000).
28. Li, H., Chaney, S., Forte, M. & Hirsh, J. Ectopic G-protein expression in dopamine and serotonin neurons blocks cocaine sensitization in *Drosophila melanogaster*. *Curr. Biol.* (in the press).
29. Robinow, S. & White, L. The locus *elav* of *Drosophila melanogaster* is expressed in all neurons at all developmental stages. *Dev. Biol.* **126**, 294–303 (1988).
30. Ellis M., O'Neill, E. & Rubin, G. Expression of *Drosophila* glass protein and evidence for negative regulation of its activity in non-neuronal cells by another DNA-binding protein. *Development* **119**, 855–865 (1993).

Acknowledgements

We thank J. Hirsh, J. Hardy, M. Farrer and H. Orr for flies and DNAs; J. Hirsh, D. Dickson, M. Froesch, K. Buckley, W. Quinn and D. Morisato for discussions; and H. Shing, L. Trakimas, A. Merola, C. Ridolfi and M. Ericsson for technical assistance. M.B.F. thanks J. Gusella and the American Parkinson Disease Foundation for encouragement. Support was provided by a Howard Hughes Physician Postdoctoral Fellowship and a grant from the N.I.A. to M.B.F. and by a grant from the N.I.H. to W.B.

Correspondence and requests for materials should be addressed to M.B.F. (e-mail: mel_feany@hms.harvard.edu).

Interleukin-1 polymorphisms associated with increased risk of gastric cancer

Emad M. El-Omar[†], Mary Carrington[‡], Wong-Ho Chow^{*}, Kenneth E. L. McColl[§], Jay H. Bream^{||}, Howard A. Young^{||}, Jesus Herrera[‡], Jolanta Lissowska[¶], Chiu-Chin Yuan[‡], Nathaniel Rothman^{*}, George Lanyon[§], Maureen Martin[‡], Joseph F. Fraumeni Jr^{*} & Charles S. Rabkin^{*}

^{*} Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, Maryland, USA

[†] Department of Medicine and Therapeutics, Aberdeen University, Aberdeen, UK
[‡] Intramural Research Support Program, Science Applications International Corporation Frederick, National Cancer Institute-Frederick Cancer Research and Development Centre, Maryland, USA

[§] Department of Medicine and Therapeutics, Western Infirmary, Glasgow, UK
^{||} Division of Basic Sciences, National Cancer Institute, Frederick Cancer Research and Development Centre, Maryland, USA

[¶] Division of Cancer Epidemiology and Prevention, Cancer Centre and M. Skłodowska-Curie Institute of Oncology, Warsaw, Poland

***Helicobacter pylori* infection is associated with a variety of clinical outcomes including gastric cancer and duodenal ulcer disease¹. The reasons for this variation are not clear, but the gastric physiological response is influenced by the severity and anatomical distribution of gastritis induced by *H. pylori*. Thus, individuals**

with gastritis predominantly localized to the antrum retain normal (or even high) acid secretion², whereas individuals with extensive corpus gastritis develop hypochlorhydria and gastric atrophy³, which are presumptive precursors of gastric cancer⁴. Here we report that interleukin-1 gene cluster polymorphisms suspected of enhancing production of interleukin-1-beta are associated with an increased risk of both hypochlorhydria induced by *H. pylori* and gastric cancer. Two of these polymorphism are in near-complete linkage disequilibrium and one is a TATA-box polymorphism that markedly affects DNA-protein interactions *in vitro*. The association with disease may be explained by the biological properties of interleukin-1-beta, which is an important pro-inflammatory cytokine⁵ and a powerful inhibitor of gastric acid secretion^{6,7}. Host genetic factors that affect interleukin-1-beta may determine why some individuals infected with *H. pylori* develop gastric cancer while others do not.

H. pylori infects half of the world's population and has been implicated in the pathogenesis of gastric cancer¹, the second-most common malignancy worldwide⁸. The mechanism of *H. pylori*-induced carcinogenesis is not clear. The infection almost always causes inflammation of the gastric mucosa, the distribution and severity of which varies widely and affects the clinical outcome. Gastritis that is confined to the antral region is associated with excessive acid production and a high risk of duodenal ulcer disease². In contrast, gastritis involving the acid-secreting corpus region leads to hypochlorhydria, progressive gastric atrophy³ and an increased risk of gastric cancer^{4,9}. These pre-cancerous changes are unfavourable to *H. pylori* growth and some gastric cancers arise long after the infection has disappeared. Duodenal ulceration and gastric cancer seem to be mutually exclusive outcomes of *H. pylori* infection that cannot be explained by differences in bacterial virulence factors alone, as virulent strains seem to be equally associated with both conditions^{10,11}. An alternative explanation is that host genetic factors (in conjunction with bacterial and/or environmental factors) determine the immune and inflammatory responses to *H. pylori* infection. A critical factor in *H. pylori*-induced gastric carcinogenesis is gastric acid secretion, which both influences and is influenced by *H. pylori*-induced gastritis. If inflammation of the corpus mucosa is severe, acid secretion is inhibited and eventually lost through the destruction of gastric glands³. Furthermore, pharmacological inhibition of acid secretion leads to the re-distribution of *H. pylori*-induced gastritis with a reduced intensity of antral inflammation

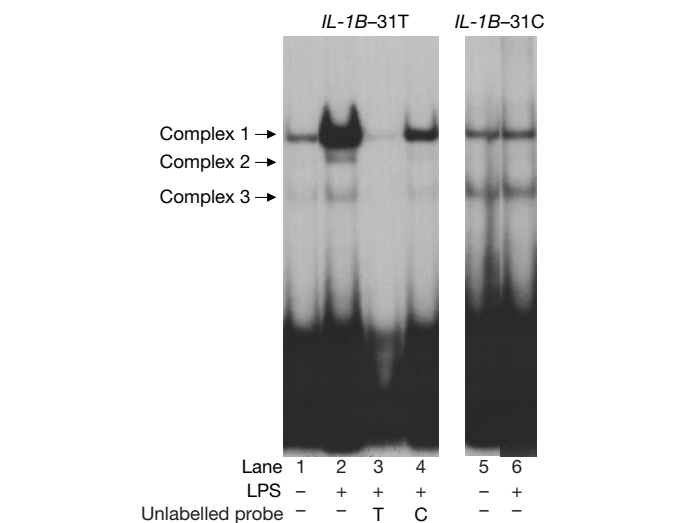


Figure 1 Differential binding patterns between the T- and C-bearing alleles of the *IL-1B* promoter. Nuclear extracts from fresh human monocytes stimulated for 30 min with lipopolysaccharide (LPS) (+) show evidence of induction of DNA-protein complexes 1 and 2 on the *IL-1B* - 31T oligonucleotide. Complex 3 and free probe were equivalent across all lanes. Cold competition with a 100-fold excess of unlabelled probe (T) blocked the formation of complex 1 (lane 3), whereas cross-competition with unlabelled *IL-1B* - 31C (C) only partially blocked complex 1 (lane 4). Results were similar with monocyte extracts from three different donors.

and an increased intensity in the corpus, and may progress to gastric atrophy in the long term¹². Endogenous factors that suppress acid secretion could also contribute to corpus gastritis and atrophy.

The interleukin-1 (*IL-1*) gene cluster on chromosome 2q contains 3 related genes within a 430-kilobase (kb) region, *IL-1A*, *IL-1B* and *IL-1RN*, which encode the pro-inflammatory cytokines IL-1 α and IL-1 β as well as their endogenous receptor antagonist IL-1ra respectively⁵. IL-1 β is upregulated in the presence of *H. pylori* and is important in initiating and amplifying the inflammatory response to this infection¹³⁻¹⁵. IL-1 β is also a potent inhibitor of gastric acid secretion^{6,7}; on a molar basis it is estimated to be 100-fold more potent than proton pump inhibitors and 6,000-fold more potent

Table 1 Estimated haplotype frequencies and linkage disequilibrium coefficients

Population	Loci	Haplotype*						Disequilibrium†		
		1-1	1-2	2-1	2-2	1-3,4,5	2-3,4,5	D'	χ^2	p
GCR controls (n = 100)	<i>IL-1B</i> - 31/ <i>IL-1B</i> - 511	0.610	0.010	0.005	0.375			0.98	95.2	0.0001
	<i>IL-1B</i> - 31/ <i>IL-1B</i> + 3954	0.420	0.200	0.380	0.000			-0.53	27.8	0.0001
	<i>IL-1B</i> - 31/ <i>IL-1RN</i>	0.519	0.096	0.131	0.244	0.005	0.005	0.52	26.6	0.0001
	<i>IL-1B</i> + 3954/ <i>IL-1RN</i>	0.495	0.295	0.155	0.045	0.010	0.000	-0.20	3.8	0.05
Low acid GCR (n = 45)	<i>IL-1B</i> - 31/ <i>IL-1B</i> - 511	0.422	0.000	0.022	0.556			0.96	41.4	0.0001
	<i>IL-1B</i> - 31/ <i>IL-1B</i> + 3954	0.273	0.149	0.560	0.018			-0.27	3.3	0.07
	<i>IL-1B</i> - 31/ <i>IL-1RN</i>	0.330	0.092	0.203	0.375	0.000	0.000	0.48	10.5	0.001
	<i>IL-1B</i> + 3954/ <i>IL-1RN</i>	0.408	0.425	0.125	0.042	0.000	0.000	-0.17	1.4	0.24
Normal acid GCR (n = 58)	<i>IL-1B</i> - 31/ <i>IL-1B</i> - 511	0.681	0.017	0.000	0.302			0.95	52.1	0.0001
	<i>IL-1B</i> - 31/ <i>IL-1B</i> + 3954	0.527	0.171	0.275	0.027			-0.20	2.3	0.13
	<i>IL-1B</i> - 31/ <i>IL-1RN</i>	0.611	0.087	0.130	0.155	0.000	0.017	0.47	12.6	0.0004
	<i>IL-1B</i> + 3954/ <i>IL-1RN</i>	0.561	0.223	0.180	0.018	0.017	0.000	-0.25	3.6	0.06
Gastric cancer controls (n = 429)	<i>IL-1B</i> - 31/ <i>IL-1B</i> - 511	0.699	0.002	0.000	0.298			1.00	426.6	0.0001
	<i>IL-1B</i> - 31/ <i>IL-1B</i> + 3954	0.498	0.203	0.250	0.048			-0.16	11.2	0.0008
	<i>IL-1B</i> - 31/ <i>IL-1RN</i>	0.612	0.084	0.109	0.185	0.005	0.004	0.51	112.6	0.0001
	<i>IL-1B</i> + 3954/ <i>IL-1RN</i>	0.525	0.214	0.196	0.056	0.009	0.000	-0.07	2.1	0.14
Gastric cancer cases (n = 366)	<i>IL-1B</i> - 31/ <i>IL-1B</i> - 511	0.576	0.008	0.003	0.413			0.98	352.4	0.0001
	<i>IL-1B</i> - 31/ <i>IL-1B</i> + 3954	0.418	0.166	0.352	0.063			-0.19	12.8	0.0003
	<i>IL-1B</i> - 31/ <i>IL-1RN</i>	0.423	0.155	0.152	0.259	0.007	0.004	0.39	54.7	0.0001
	<i>IL-1B</i> + 3954/ <i>IL-1RN</i>	0.413	0.350	0.162	0.064	0.007	0.004	-0.17	11.0	0.0009

Data given for pairs of *IL-1* loci in gastric cancer relatives (GCR), gastric cancer cases and respective control populations.

* For all *IL-1B* loci, C is denoted by allele 1 and T is denoted by allele 2.

† D' = D/D_{max} for association of most common alleles at each locus.

than H_2 antagonists¹⁶. Three diallelic polymorphisms in *IL-1B* have been reported, all representing C–T base transitions, at positions –511, –31 and +3954 base pairs (bp) from the transcriptional start site¹⁷. There are conflicting data regarding the functional effects of these polymorphisms on IL-1 β production^{18,19}. The *IL-1RN* gene has a penta-allelic 86-bp tandem repeat (VNTR) in intron 2, of which the less common allele 2 (*IL-1RN**2) is associated with a wide range of chronic inflammatory and autoimmune conditions¹⁷. *IL-1RN**2 is associated with enhanced IL-1 β production *in vitro*¹⁹, but data regarding its effects on IL-1 α production are contradictory^{20–22}.

To determine whether these polymorphisms are important with respect to the different outcomes of *H. pylori* infections, we studied their effects on gastric physiology in healthy subjects. We had previously shown that a cohort of 149 first-degree relatives of gastric cancer patients from the West of Scotland had a high prevalence of hypochlorhydria (defined as a pentagastrin-stimulated peak acid output of less than 15 mmol h^{–1}) in association with *H. pylori* infection²³. Of the 103 (69%) gastric cancer relatives (GCR) infected with the *H. pylori*, 45 had hypochlorhydria and gastric atrophy and 58 had normal or high acid secretion. One hundred unselected newborns from the West of Scotland were available as population controls. We also had available a case-control study of gastric cancer, based on a population from Warsaw, Poland, in which there were 393 gastric cancer cases and 430 controls with DNA samples²⁴. We were thus able to determine whether genotypes that modify IL-1 β are associated with low acid secretion and gastric atrophy, and whether these same genotypes increase the risk of gastric cancer. These studies were reviewed and approved by the Institutional Review Boards and Ethics Committees of the US National Cancer Institute, the University of Glasgow Hospitals NHS Trust, and the M. Skłodowska-Curie Memorial Cancer Centre (Warsaw).

In both the Scottish and Polish control populations, the alleles at the individual loci of *IL-1B* and *IL-1RN* were in Hardy–Weinberg equilibrium, with non-significant χ^2 values. There was marked linkage disequilibrium among the loci within the gene cluster (see also ref. 25). Linkage disequilibrium between *IL-1B* – 31 and *IL-1B* – 511 was almost total, with 99.5% of the inferred haplotypes (*IL-1B* – 31/*IL-1B* – 511) in the combined control groups consisting of either T–T or C–C (Table 1). There was also strong linkage disequilibrium ($D' \approx 0.5$) between *IL-1B* – 31 and *IL-1RN* in both populations. However, the two groups differed with respect to linkage disequilibrium between *IL-1B* – 31 and *IL-1B* + 3954, which was strong in the Scottish population controls but relatively weak ($D' = -0.16$) in the Polish controls. There was no significant linkage disequilibrium between *IL-1B* + 3954 and *IL-1RN* in either control group.

There were no significant differences in genotype frequency for any *IL-1* marker between the GCR and controls or between the total

H. pylori-infected and uninfected GCR (Table 2). Nevertheless, among the infected GCR, those with low acid secretion had a significantly higher frequency of the pro-inflammatory *IL-1RN**2 allele and the T–T haplotype of *IL-1B* – 31 and *IL-1B* – 511 (*IL-1B* – 31T/*IL-1B* – 511T), as compared with the GCR with normal or high acid secretion. Carriers of the T allele of *IL-1B* – 31 (*IL-1B* – 31T+) had an age-adjusted odds ratio of 9.1 (95% confidence interval (CI), 2.2–37), and there was little difference between the homozygous and heterozygous carriers (Table 2). *IL-1RN**2 homozygotes (*IL-1RN**2/*2) were also at increased risk of hypochlorhydria, although risk among the *IL-1RN**2 heterozygotes was not significantly increased (Table 2). In a logistic regression model including both factors, the estimated age-adjusted odds ratios for *IL-1B* – 31T+ and *IL-1RN**2/*2 were 7.5 (95% CI, 1.8–31) and 2.1 (95% CI, 0.7–6.3), respectively. The *IL-1B* + 3954 genotype was not associated with the risk of hypochlorhydria.

There were similar associations between these alleles and the risk of gastric cancer. Carriers of *IL-1B* – 31T had an increased gastric cancer risk at an odds ratio of 1.9 (95% CI, 1.5–2.6), with no significant difference between homozygotes and heterozygotes (Table 3). Moreover, *IL-1RN**2 was associated with an increased risk in homozygotes but not in heterozygotes (Table 3). In a logistic regression model including both of these genotypes, the estimated odds ratios for *IL-1B* – 31T+ and *IL-1RN**2/*2 were 1.6 (95% CI, 1.2–2.2) and 2.9 (95% CI, 1.9–4.4), respectively. *IL-1B* + 3954T homozygotes seemed to be protected against gastric cancer, although the effect did not reach statistical significance (Table 3).

The estimated effects of *IL-1B* – 31T+ and *IL-1RN**2/*2 were similar in subgroups of gastric cancer cases defined by age, sex, histological type and anatomical site (data not shown). Furthermore, adjustment for other reported risk factors for gastric cancer, including tobacco and alcohol use, ABO blood group and family history of gastric cancer, did not substantially alter the estimates (data not shown).

The *IL-1B* – 31T/*IL-1RN**2 haplotype imparted a greatly increased risk of gastric cancer, as compared with having no copy of *IL-1B* – 31T and at least one copy of the other *IL-1RN* alleles. Twenty-two per cent of cases (compared with eight per cent of controls) had this haplotype, either in its homozygous form or with *IL-1B* – 31C/*IL-1RN**2, with an odds ratio of 4.4 (95% CI, 2.8–6.9). In the absence of *IL-1B* – 31T, homozygous *IL-1RN**2 was associated with a similarly elevated odds ratio of 5.3 (95% CI, 1.9–14), although this genotype combination was uncommon because of linkage disequilibrium between the two loci. In contrast, the odds ratio for *IL-1B* – 31T+ with no more than one copy of *IL-1RN**2 was only 1.7 (95% CI, 1.2–2.3). Nonetheless, because its effect is observed in both homozygotes and heterozygotes, the *IL-1B* – 31T allele accounts for a greater proportion of excess gastric cancer cases than the *IL-1RN**2 allele, despite their comparable frequencies. The

Table 2 *IL-1* genotype frequencies in gastric cancer relatives (GCR) and controls

Locus	Genotype	<i>H. pylori</i> -infected GCR			Uninfected GCR (<i>n</i> = 46)	Population controls (<i>n</i> = 100)
		Low acid (<i>n</i> = 45)	Normal acid (<i>n</i> = 58)	Odds ratio (95% CI)*		
<i>IL-1B</i> – 31	C/C	5	30	1.0	23	37
	C/T	28	21	8.1 (2.0–33)	18	50
	T/T	12	7	13.6 (2.6–71)	5	13
<i>IL-1B</i> – 511	C/C	5	29	1.0	23	36
	C/T	30	21	8.3 (2.0–34)	18	51
	T/T	10	8	11.4 (2.2–58)	5	13
<i>IL-1B</i> + 3954	C/C	30	35	1.0	28	67
	C/T	15	23	0.8 (0.3–1.9)	15	26
	T/T	0	0		3	7
<i>IL-1RN</i>	1/1	17	35	1.0	24	42
	1/2	14	14	2.4 (0.9–6.2)	15	44
	1/3, 4, 5	0	2	0	0	2
	2/2	14	7	5.6 (1.8–17)	7	12

* Odds ratio for low acid versus normal/high acid, adjusted for age and within-family sampling. CI, confidence interval.

fraction of gastric cancer in the population that is attributable to *IL-1B* - 31T is estimated to be 31%, compared with 18% for *IL-1RN**2. Their combined population attributable fraction is estimated to be 38%, which represents the fraction of gastric cancer cases that are caused by these *IL-1* alleles.

The *IL-1B* - 31 polymorphism involves a TATA sequence in the *IL-1B* promoter. To investigate the effect of *IL-1B* - 31 variants on the induction of *IL-1β*, we used electrophoretic mobility-shift analysis to assess their DNA-binding activity *in vitro*. Synthetic allele-specific oligonucleotides representing the polymorphic *IL-1B* - 31 sites were incubated with nuclear protein extracts from non-stimulated human monocytes and monocytes activated by lipopolysaccharide (LPS). LPS stimulation induced a fivefold increase in DNA binding (complex 1) on the *IL-1B* - 31T oligonucleotide (Fig. 1, lanes 1 and 2). In contrast, LPS failed to induce complex 1 formation on the *IL-1B* - 31C oligonucleotide (Fig. 1, lanes 5 and 6). Furthermore, complex 1 formation on radiolabelled *IL-1B* - 31T was specifically blocked by competition with the unlabelled *IL-1B* - 31T, but not the *IL-1B* - 31C, oligonucleotide (Fig. 1, lanes 3 and 4). These results indicate that one or more proteins (presumably transcription factors) in complex 1 may be unable to interact with the C-bearing *IL-1B* - 31 allele to form the transcription initiation complex. In parallel experiments assessing allele-specific oligonucleotides for *IL-1B* - 511, there were no differences in binding activity (data not shown), indicating that the effect of *IL-1B* - 511 may be mediated by linkage disequilibrium with the TATA box polymorphism. On the basis of these results and previously published functional data for *IL-1RN*, hypochlorhydria and gastric cancer may be associated with alleles of *IL-1B* - 31 and *IL-1RN* that enhance *IL-1β* production.

Here we demonstrate that pro-inflammatory genotypes of the *IL-1* loci (*IL-1B* - 31T+ and *IL-1RN**2/*2) increase both the likelihood of a chronic hypochlorhydric response to *H. pylori* infection and the risk of gastric cancer, presumably by altering *IL-1β* levels in the stomach. While the pro-inflammatory effects of high *IL-1β* concentrations may facilitate the clearance of *H. pylori* from the gastric mucosa, the concomitant inhibition of acid secretion allows the spread of *H. pylori*-induced inflammation from the antrum to the corpus. This functional inhibition is initially reversible but the progressive destruction of parietal cells eventually leads to irreversible hypochlorhydria. A decreased flow of gastric secretions may therefore heighten mucosal damage by allowing the accumulation of bacterial toxins and by-products of inflammation that would normally be diluted and flushed out. The reactive oxygen and nitrogen oxide species that are derived from inflammation are known mutagens, while hypochlorhydria permits superinfection by other bacteria that enhance the production of highly carcinogenic *N*-nitroso compounds²⁶.

H. pylori-induced hypochlorhydria also markedly reduces the levels of vitamin C in gastric juice²⁷, further facilitating the forma-

tion of *N*-nitroso compounds. The opportunities for DNA damage caused by this cascade of genotoxic factors are amplified by the increased rate of cell turnover in inflamed mucosa. Therefore, genotypes that enhance *IL-1β* production may favour the initiation of a set of responses to *H. pylori* that result in hypochlorhydria, corpus atrophy and an increased risk of gastric cancer.

Our findings complement the most widely accepted multi-stage model of gastric carcinogenesis⁴, and provide insights into the etiological role of *H. pylori*. We have shown at least two susceptibility loci in the *IL-1* gene cluster for gastric cancer and its precursors. The effects of these loci influence an early stage of the disease process and require the presence of *H. pylori* infection. Progression towards cancer is probably influenced by other components of the host genetic constitution acting epistatically, as well as by dietary and other factors in the environment. *IL-1B* - 31T/*IL-1RN**2 constitutes a pro-inflammatory haplotype that is in strong linkage disequilibrium, at least in Caucasian populations. We speculate that this linkage disequilibrium reflects past selective pressures for or against enhancement of the *IL-1β* response to environmental challenges. Whereas pro-inflammatory genotypes may be advantageous for the host response to some infections, the vigorous *IL-1β* production associated with *H. pylori* gastric infection may exacerbate mucosal damage and increase the risk of eventual neoplasia. □

Methods

IL-1 genotyping

IL-1B polymorphisms were distinguished by two separate methods, polymerase chain reaction single-strand conformation polymorphism (PCR-SSCP) and 5' nuclease PCR assays (TaqMan). For PCR-SSCP, 50 ng DNA was amplified in a GeneAmp PCR System 9700 (PE Applied Biosystems), using the primer pairs listed in the Supplementary Information. Amplification was performed in a volume of 20 μl, containing 10 mM Tris-HCl pH 8.3, 50 mM KCl, 1.5 mM MgCl₂, 200 μM each of dATP, dTTP and dGTP, 100 μM dCTP, 1 μCi of [α-³²P]dCTP (3,000 Ci mol⁻¹), 80 ng of each primer and 1 unit of Taq polymerase. The thermocycling conditions were as follows: 94 °C for 10 min; then 5 cycles of 94 °C for 30 s, 65 °C for 30 s and 72 °C for 30 s; then 30 cycles of 94 °C for 30 s, 60 °C for 30 s and 72 °C for 30 s; then 5 cycles of 94 °C for 30 s, 55 °C for 30 s and 72 °C for 30 s. SSCP analysis of the radiolabelled amplification products was performed as described²⁸.

For TaqMan assays, primers (Operon Technologies Inc.) and probes (PE Applied Biosystems) were designed using Primer Express software (PE Applied Biosystems). Probes for the T or C allele were 5'-labelled with either FAM (6-carboxyfluorescein) or VIC fluorogenic dyes, and 3'-labelled with TAMRA (6-carboxytetramethylrhodamine) quencher. PCR amplification was performed in a volume of 25 μl containing 50 ng genomic DNA, 1× TaqMan Universal Master Mix (PE Applied Biosystems), 200 nM for each probe and 900 nM for primers. Cycling conditions were 50 °C for 2 min, 95 °C for 10 min, then 40 cycles of 95 °C for 15 s and 62 °C for 1 min, as recommended by the manufacturer. Thermal cycling of optical plates was performed in GeneAmp PCR System 9700 and endpoint analysis was performed in the ABI PRISM 7700 Sequence Detection System (PE Applied Biosystems). Sequences of primers and probes for *IL-1B* TaqMan assays are available from the authors.

For *IL-1RN*, genomic DNA was amplified using PCR under the conditions described above for *IL-1B*, using forward primer 5'-CCCTCAGCAACACTCC-3' and reverse primer 5'-GGTCAGAAGGCGCAGAGA-3'. The PCR products were separated by electrophoresis on 2% agarose gels and stained with ethidium bromide. Alleles were sized relative to a 1-kb DNA ladder and coded conventionally as follows: allele 1 = 4 repeats, allele 2 = 2 repeats, allele 3 = 5 repeats, allele 4 = 3 repeats, allele 5 = 6 repeats; the rarer alleles 3, 4 and 5 were grouped in the statistical analysis.

A total of 93% (366/393) of the gastric cancer cases, 99% (429/430) of their population controls, and all GCR subjects and their newborn population controls were successfully genotyped for all four loci.

Electrophoretic mobility shift assay (EMSA)

Nuclear extracts were prepared from freshly isolated human monocytes as described²⁹, after no stimulation or addition of 1 μg ml⁻¹ LPS derived from *E. coli* (Sigma) for 30 min. Complementary single-stranded oligonucleotides were synthesized (Life Technologies) as follows (variant nucleotides in bold):

IL-1B - 31 : 5'-TGCTTTTGAAGCC/TATAAAACAGCG-3'

IL-1B - 511 : 5'-TGACAGAGAGCTCCC/TGAGGCAGAGAAC-3'

Complementary strands were annealed by combining 2 μg of each oligonucleotide and 6 μl of 10× annealing buffer (500 mM Tris, 100 mM MgCl₂ and 50 mM dithiothreitol) in a 60-μl reaction, placing in a boiling water bath for 5 min and allowing to cool to room temperature. The DNA-protein binding reaction was conducted in a 20-μl volume containing 7 μg of nuclear protein extract, 1 μg poly (dI-dC) (Sigma), 4 μl of 5× binding

Table 3 *IL-1* genotype frequencies in gastric cancer cases and controls

Locus	Genotype	Cases (n = 366)	Controls (n = 429)	Odds ratio (95% CI)
<i>IL-1B</i> - 31	C/C	128	219	1.0
	C/T	172	164	1.8 (1.3-2.4)
	T/T	66	46	2.5 (1.6-3.8)
<i>IL-1B</i> - 511	C/C	127	217	1.0
	C/T	170	166	1.8 (1.3-2.4)
	T/T	69	46	2.6 (1.7-3.9)
<i>IL-1B</i> + 3954	C/C	212	242	1.0
	C/T	140	158	1.0 (0.8-1.4)
	T/T	14	29	0.6 (0.3-1.1)
<i>IL-1RN</i>	1/1	148	230	1.0
	1/2	117	152	1.2 (0.9-1.6)
	1/3, 4, 5	8	7	1.8 (0.7-4.8)
	2/2	93	39	3.7 (2.4-5.7)
	2/5	0	1	0

buffer (60 mM HEPES, 7.5 mM MgCl₂, 300 mM KCl, 1 mM ethylenediamine-tetraacetic acid, 2.5 mM dithiothreitol, 50% glycerol and 4-(2-aminoethyl)-benzenesulphonyl fluoride hydrochloride) and 2.5 × 10⁴ c.p.m. of ³²P-labelled oligonucleotide probe. UN-SCAN-IT 5.1 software (Silk Scientific) was used for densitometric analysis of the autoradiographs.

Statistical analysis

Hardy–Weinberg equilibrium of alleles at individual loci was assessed by χ^2 statistics. Haplotype frequencies for pairs of alleles were estimated using the Estimating Haplotype-frequencies (EH) software program (<http://linkage.rockefeller.edu/software/eh>). Linkage disequilibrium coefficients $D' = D/D_{\max}$ and χ^2 values were calculated for pairs of the most common alleles at each locus using the LINKDOS software program (distributed with GENEPOP, <http://ftp.cfe.cnrs-mop.fr/pub/PC/MSDOS/GENEPOP/>). Odds ratios with Cornfield 95% confidence intervals and logistic regression models controlling for the effects of possible confounders were computed using STATA version 5.0 software (STATA Press). The odds ratios for hypochlorhydria were age-adjusted (categorized as ≤35, 36–45, 46–55 and >55 years) because of its age-dependence²² and their confidence intervals were based on robust variance estimates³⁰ which adjust for within-family correlation, to account for sampling of several members of a given family.

Received 21 October 1999; accepted 4 February 2000.

- Blaser, M. J. *Helicobacter pylori* and gastric diseases. *Br. Med. J.* **316**, 1507–1510 (1998).
- El-Omar, E. et al. *Helicobacter pylori* infection and abnormalities of acid secretion in patients with duodenal ulcer disease. *Gastroenterology* **109**, 681–691 (1995).
- El-Omar, E. et al. *Helicobacter pylori* infection and chronic gastric acid hyposecretion. *Gastroenterology* **113**, 15–24 (1997).
- Correa, P. Human gastric carcinogenesis: a multistep and multifactorial process—First American Cancer Society Award Lecture on Cancer Epidemiology and Prevention. *Cancer Res.* **52**, 6735–6740 (1992).
- Dinarello, C. A. Biologic basis for interleukin-1 in disease. *Blood* **87**, 2095–2147 (1996).
- Wallace, J. L., Cucala, M., Mugridge, K. & Parente, L. Secretagogue-specific effects of interleukin-1 on gastric acid secretion. *Am. J. Physiol.* **261**, G559–G564 (1991).
- Beales, I. L. & Calam, J. Interleukin 1 beta and tumour necrosis factor alpha inhibit acid secretion in cultured rabbit parietal cells by multiple pathways. *Gut* **42**, 227–234 (1998).
- Parkin, D. M., Pisani, P. & Ferlay, J. Estimates of the worldwide incidence of 25 major cancers in 1990. *Int. J. Cancer* **80**, 827–841 (1999).
- Sipponen, P. Gastric cancer—a long-term consequence of *Helicobacter pylori* infection? *Scand. J. Gastroenterol.* **201**, (suppl.), 24–27 (1994).
- Yamaoka, Y. et al. Relationship between *Helicobacter pylori* iceA, cagA, and vacA status and clinical outcome: studies in four different countries. *J. Clin. Microbiol.* **37**, 2274–2279 (1999).
- Hansson, L. E. et al. The risk of stomach cancer in patients with gastric or duodenal ulcer disease. *N. Engl. J. Med.* **335**, 242–249 (1996).
- Kuipers, E. J. et al. Atrophic gastritis and *Helicobacter pylori* infection in patients with reflux esophagitis treated with omeprazole or fundoplication. *N. Engl. J. Med.* **334**, 1018–1022 (1996).
- Noach, L. A. et al. Mucosal tumor necrosis factor-alpha, interleukin-1 beta, and interleukin-8 production in patients with *Helicobacter pylori* infection. *Scand. J. Gastroenterol.* **29**, 425–429 (1994).
- Basso, D. et al. *Helicobacter pylori* infection enhances mucosal interleukin-1 beta, interleukin-6, and the soluble receptor of interleukin-2. *Int. J. Clin. Lab. Res.* **26**, 207–210 (1996).
- Jung, H. C., Kim, J. M., Song, I. S. & Kim, C. Y. *Helicobacter pylori* induces an array of pro-inflammatory cytokines in human gastric epithelial cells: quantification of mRNA for interleukin-8, -1 alpha/beta, granulocyte-macrophage colony-stimulating factor, monocyte chemoattractant protein-1 and tumour necrosis factor-alpha. *J. Gastroenterol. Hepatol.* **12**, 473–480 (1997).
- Wolfe, M. M. & Nompreggi, D. J. Cytokine inhibition of gastric acid secretion—a little goes a long way. *Gastroenterology* **102**, 2177–2178 (1992).
- Bidwell, J. L. et al. Cytokine gene polymorphism in human disease: on-line databases. (<http://www.pam.bris.ac.uk/services/GAI/cytokine4.htm>).
- Pociot, F. et al. A TagI polymorphism in the human interleukin-1 beta (IL-1 beta) gene correlates with IL-1 beta secretion in vitro. *Eur. J. Clin. Invest.* **22**, 396–402 (1992).
- Santtila, S., Savinainen, K. & Hurme, M. Presence of the IL-1RA allele 2 (IL1RN*2) is associated with enhanced IL-1beta production in vitro. *Scand. J. Immunol.* **47**, 195–198 (1998).
- Andus, T. et al. Imbalance of the interleukin 1 system in colonic mucosa—association with intestinal inflammation and interleukin 1 receptor antagonist genotype 2. *Gut* **41**, 651–657 (1997).
- Danis, V. A., Millington, M., Hyland, V. J. & Grennan, D. Cytokine production by normal human monocytes: inter-subject variation and relationship to an IL-1 receptor antagonist (IL-1RA) gene polymorphism. *Clin. Exp. Immunol.* **99**, 303–310 (1995).
- Tountas, N. A. et al. Functional and ethnic association of allele 2 of the interleukin-1 receptor antagonist gene in ulcerative colitis. *Gastroenterology* **117**, 806–813 (1999).
- El-Omar, E. M. et al. Increased prevalence of precancerous changes in relatives of gastric cancer patients: critical role of *H. pylori*. *Gastroenterology* **118**, 22–30 (2000).
- Chow, H. W. et al. Risk of stomach cancer in relation to consumption of cigarettes, alcohol, tea and coffee in Warsaw, Poland. *Int. J. Cancer* **81**, 871–876 (1999).
- Cox, A., Camp, N. J., Nicklin, M. J., Di Giovine, F. S. & Duff, G. W. An analysis of linkage disequilibrium in the interleukin-1 gene cluster, using a novel grouping method for multiallelic markers. *Am. J. Hum. Genet.* **62**, 1180–1188 (1998).
- Stockbruegger, R. W. et al. Pernicious anaemia, intragastric bacterial overgrowth, and possible consequences. *Scand. J. Gastroenterol.* **19**, 355–364 (1984).
- Ruiz, B. et al. Vitamin C concentration in gastric juice before and after anti-*Helicobacter pylori* treatment. *Am. J. Gastroenterol.* **89**, 533–539 (1994).
- Carrington, M. et al. Typing of HLA-DQA1 and DQB1 using DNA single-strand conformation polymorphism. *Hum. Immunol.* **33**, 208–212 (1992).
- Yu, C. R. et al. Differential utilization of Janus kinase-signal transducer activator of transcription signaling pathways in the stimulation of human natural killer cells by IL-2, IL-12, and IFN-alpha. *J. Immunol.* **157**, 126–137 (1996).

30. Royall, R. M. Model robust confidence intervals using maximum likelihood estimators. *Int. Stat. Rev.* **54**, 221–226 (1986).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

Acknowledgements

We thank A. Goldstein and N. Chatterjee for advice on genetic and statistical issues, N. Dunlap for technical assistance, D. Gillen for help with subject recruitment and J. Goedert for suggestions. E.M.E. received a European *Helicobacter pylori* Study Group Research Fellowship from the Digestive Disorders Foundation, UK. This project was partly funded by the National Cancer Institute, National Institutes of Health, USA.

Correspondence and requests for materials should be addressed to E.M.E. (e-mail: elomare@mail.nih.gov).

PKC-θ is required for TCR-induced NF-κB activation in mature but not immature T lymphocytes

Zuoming Sun*, Christopher W. Arendt*, Wilfried Ellmeier*†, Edward M. Schaeffer‡§, Mary Jean Sunshine*†, Leena Gandhi*, Justin Annes*, Daniela Petrzilka*†, Abraham Kupfer||, Pamela L. Schwartzberg‡ & Dan R. Littman*†

† Howard Hughes Medical Institute and * Molecular Pathogenesis Program, Skirball Institute of Biomolecular Medicine, New York University School of Medicine, New York, New York 10016, USA

‡ National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland 20892, USA

§ Department of Pathology, University of Chicago, Chicago, Illinois 60637, USA

|| Division of Basic Science, Department of Pediatrics, National Jewish Medical and Research Center, Denver, Colorado 80262, USA

Productive interaction of a T lymphocyte with an antigen-presenting cell results in the clustering of the T-cell antigen receptor (TCR) and the recruitment of a large signalling complex to the site of cell–cell contact^{1,2}. Subsequent signal transduction resulting in cytokine gene expression requires the activation of one or more of the multiple isoenzymes of serine/threonine-specific protein kinase C (PKC)³. Among the several PKC isoenzymes expressed in T cells, PKC-θ is unique in being rapidly recruited to the site of TCR clustering⁴. Here we show that PKC-θ is essential for TCR-mediated T-cell activation, but is dispensable during TCR-dependent thymocyte development. TCR-initiated NF-κB activation was absent from PKC-θ^{-/-} mature T lymphocytes, but was intact in thymocytes. Activation of NF-κB by tumour-necrosis factor α and interleukin-1 was unaffected in the mutant mice. Although studies in T-cell lines had suggested that PKC-θ regulates activation of the JNK signalling pathway^{5,6}, induction of JNK was normal in T cells from mutant mice. These results indicate that PKC-θ functions in a unique pathway that links the TCR signalling complex to the activation of NF-κB in mature T lymphocytes.

We inactivated the gene encoding PKC-θ by homologous recombination in embryonic stem cells by replacing the exon encoding the ATP-binding site of the kinase domain (amino-acid residues 396–451) with the *neomycin* resistance gene. Homozygous mutant mice seemed normal and were fertile. Immunoblot analysis with antibodies directed against sequences outside the deleted coding region of PKC-θ failed to detect any protein product, even of smaller than normal size, in thymocytes or T cells from the mutant animals (data not shown). Flow cytometric analyses of cells from thymus, spleen and lymph nodes of the mutant mice were indistinguishable from

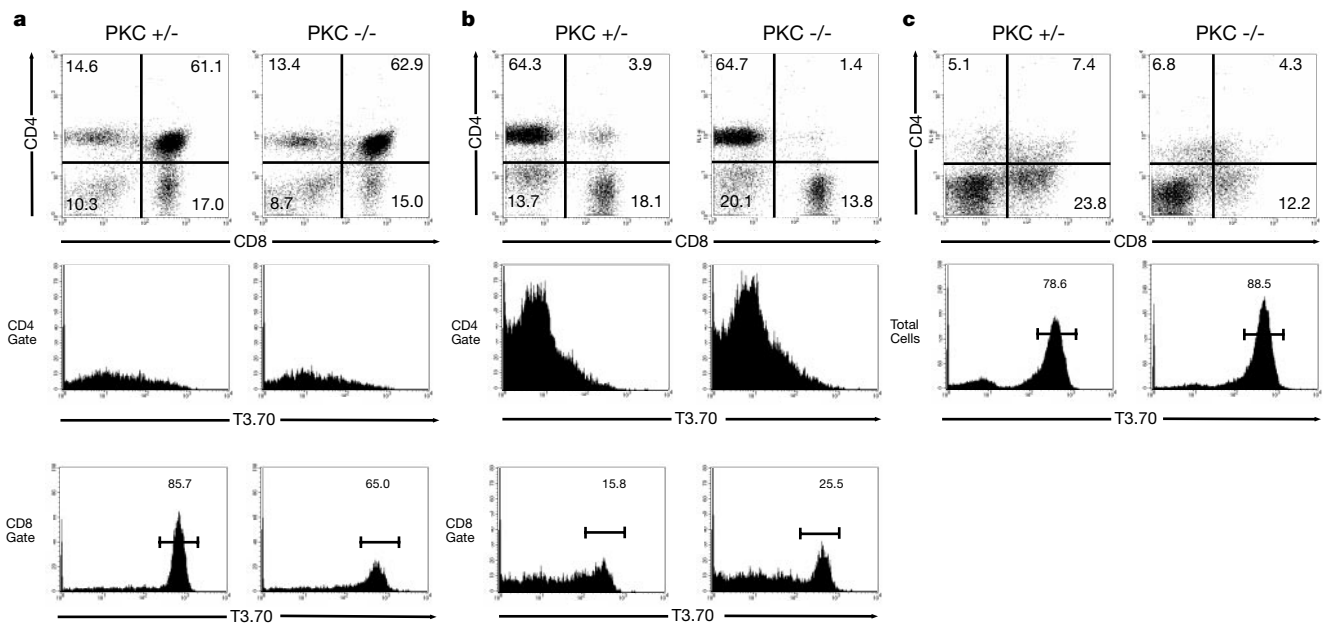


Figure 1 Thymic maturation in *PKC-θ*-null mice. **a**, Positive selection of the HY transgenic TCR, as revealed by staining with T3.70, in CD8⁺ thymocytes from female mice heterozygous or homozygous for the targeted *PKC-θ* mutation. CD4 and CD8 gates are the same as those shown in the dot plots, where the percentages of total thymocytes are indicated. **b**, Expression of the clonotypic TCR in CD8⁺, but not CD4⁺, lymph node cells

from female mice. **c**, Negative selection of double-positive thymocytes in male mice expressing the HY transgenic TCR. Total thymocyte numbers were: 1.59×10^8 and 1.13×10^8 in heterozygous and null females, respectively, and 10.2×10^6 and 8.2×10^6 in the heterozygous and null males.

those of wild-type littermate controls with regard to cell numbers and levels of surface proteins (data not shown). Further examination of thymocyte differentiation was done by mating to mice expressing a transgenic TCR specific for the male antigen, HY, in the context of the class I molecule H-2D^b (ref. 7). In female mice, T cells that express the transgenic TCR are preferentially selected to become CD8⁺ single-positive cells, and can be detected with the clonotypic antibody T3.70. In the absence of *PKC-θ*, positive selection in female animals proceeded normally (Fig. 1a, b). In male mice, negative selection, reflected as a depletion of double-positive thymocytes, was also unaffected by the loss of *PKC-θ* (Fig. 1c). The signalling function of *PKC-θ* therefore does not seem to be required for the key selection processes during thymocyte differentiation.

Because the expression of *PKC-θ* is restricted to lymphocytes of the T-cell lineage⁸, we sought to determine whether the TCR signalling pathway in peripheral T cells was affected by its absence. Proliferation of purified T lymphocytes was measured after stimulation with plate-bound anti-CD3 monoclonal antibody (mAb) with or without anti-CD28 stimulation. In response to either of these stimuli, proliferation of T cells from the homozygous mutant mice was reduced by more than 20-fold relative to that of wild-type T cells (Fig. 2a). A marked decrease in proliferation was also observed in *PKC-θ*-deficient T cells stimulated with anti-CD28 plus 12-*O*-tetradecanoylphorbol-13-acetate (TPA). Reduced proliferation of *PKC-θ*^{-/-} T cells was accompanied by a significant reduction in the level of secreted interleukin-2 (IL-2) (Fig. 2b). The addition of IL-2 at 20 units ml⁻¹ modestly restored the anti-CD3-induced proliferative response of the mutant T cells (Fig. 2a), indicating that signalling events downstream of the cytokine receptor remained intact.

Proliferation of T cells from *PKC-θ*^{-/-} mice was also substantially reduced in mixed lymphocyte reactions (Fig. 2c). To study the role of *PKC-θ* in the priming of T cells *in vivo*, we examined the responses of wild-type and mutant mice to a T-cell-dependent antigen, keyhole limpet haemocyanin (KLH). The specific response of splenic T cells to KLH was evaluated two weeks after injection, as

shown in Fig. 2d. As expected, T cells recovered from wild-type mice responded to KLH in a dose-dependent manner. However, T-cell proliferation in mutant mice remained at a baseline level identical to that in wild-type control mice injected with carrier alone. These results are consistent with a requirement for *PKC-θ* in T-dependent responses *in vivo*.

Proliferation of T cells after crosslinking of the TCR is a consequence of upregulated transcription of the *IL-2* gene and of the *IL-2Rα* chain gene, whose product contributes to a high-affinity receptor for IL-2 (ref. 9). Expression of IL-2Rα (CD25) and of CD69, a cell-surface marker on activated T cells, can also be induced by treating T cells with phorbol esters¹⁰. In response to anti-CD3/CD28 or treatment with TPA, the induction of both CD25 and CD69 was reduced or absent in the *PKC-θ*-deficient T cells (Fig. 2e). The reduced proliferative responses of T cells lacking *PKC-θ* are therefore due to decreased levels of both IL-2 and the CD25-containing high-affinity IL-2 receptor.

The induction of cytokine gene expression and T-cell proliferation after the binding of TCR and CD28 to their cognate ligands requires proximal signalling events initiated by TCR/CD3-associated protein tyrosine kinases¹¹. After anti-CD3 crosslinking, there were no apparent differences in early tyrosine phosphorylation events between T cells from mutant and control animals (data not shown). Because previous studies in cell lines implicated *PKC-θ* in the activation of various MAP kinases, particularly JNK (Jun amino-terminal kinase), we next examined these pathways^{5,6}. In *PKC-θ*-deficient T cells stimulated with anti-CD3 mAb, activation of ERK (extracellular signal-regulated kinase) exhibited a similar profile to that of wild-type cells (Fig. 3a). JNK and p38 MAP kinase (MAPK) activation in response to anti-CD3/CD28 seemed normal in mutant T cells stimulated immediately *ex vivo* (data not shown) or prestimulated with anti-CD3 to upregulate JNK protein expression¹² (Fig. 3b).

The diverse signalling pathways downstream of the TCR converge on several key transcription factors, including NF-AT, AP-1 and NF-κB, which have important roles in the expression of IL-2 and CD25 after stimulation with anti-TCR and anti-CD28 (refs 13, 14). We

therefore focused our analysis on the effect of the *PKC- θ* mutation on activation of these factors. We observed no differences in the translocation of NF-AT1 and NF-AT2 to the nucleus after the activation of mutant and wild-type T cells with anti-CD3/CD28 or with TPA plus ionomycin (Fig. 4a). In contrast, there was a marked decrease in AP-1 activation, upon stimulation either with

anti-CD3 or with TPA, in *PKC- θ* -deficient T cells in comparison with wild-type cells (Fig. 4b). AP-1 consists of Fos and Jun family members, which are transcriptionally induced early after cross-linking with anti-TCR or the treatment of T cells with phorbol esters. Because JNK activity seemed normal under conditions of TCR and phorbol stimulation that failed to activate AP-1 (Fig. 3b)

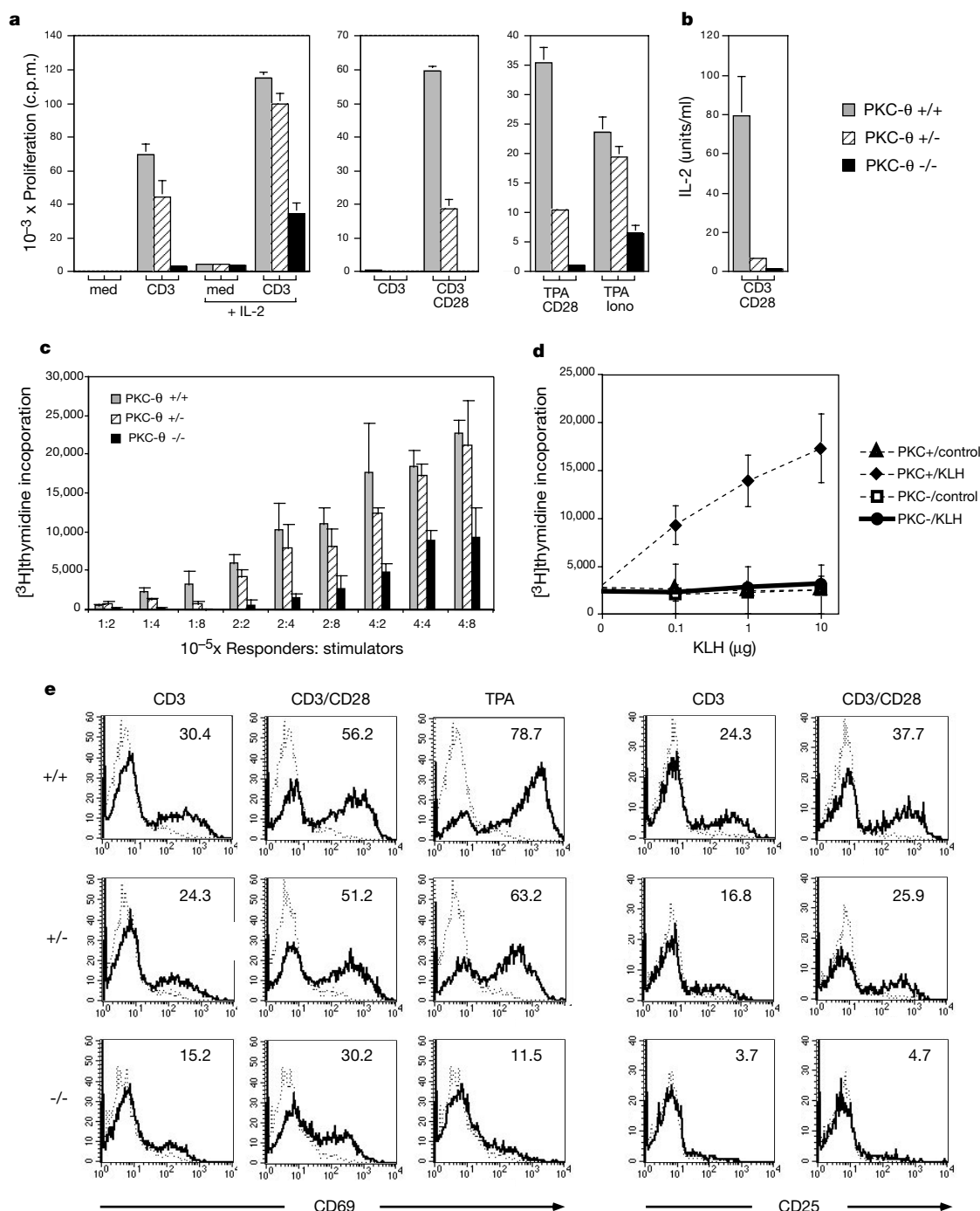


Figure 2 Impaired T-cell activation in *PKC- θ* ^{-/-} mice. **a**, Impaired proliferative responses. Left panel: purified lymph-node T cells stimulated with medium alone or plate-bound anti-CD3 ϵ (10 μ g ml⁻¹), with or without recombinant murine IL-2 (20 units ml⁻¹). Middle panel: splenic T cells stimulated with anti-CD3 ϵ (1 μ g ml⁻¹) in the presence or absence of soluble anti-CD28 (1 μ g ml⁻¹). Right panel: lymph-node T cells stimulated with TPA (10 ng ml⁻¹) and soluble anti-CD28 (1 μ g ml⁻¹) or ionomycin (Iono) (125 ng ml⁻¹). **b**, IL-2 production in supernatants of splenic T cells stimulated with plate-bound anti-CD3 ϵ (1 μ g ml⁻¹) and soluble anti-CD28 (1 μ g ml⁻¹) as measured on HT-2 responders. **c**, Mixed lymphocyte reactions of T cells from littermate mice and mitomycin C-treated BALB/c

stimulators. All proliferation assays were performed in triplicate and averaged, with error bars denoting s.d. **d**, Impaired T-dependent response to antigen in *PKC- θ* ^{-/-} mice. KLH or carrier alone was injected intraperitoneally; after 2 weeks, splenocytes were restimulated with KLH *in vitro* and T-cell proliferation was measured. Results were averaged for four mice in each group. **e**, Surface expression of CD25 and CD69, assayed by flow cytometry 14 h after stimulation with 10 μ g ml⁻¹ plate-bound anti-CD3 ϵ , with 1 μ g ml⁻¹ anti-CD3 ϵ plus 10 μ g ml⁻¹ plate-bound anti-CD28, or with 10 ng ml⁻¹ TPA. Dotted lines represent expression on unstimulated control cells.

and data not shown), the defect in AP-1 activation might reflect decreased expression of Fos and/or Jun in the absence of PKC- θ .

NF- κ B/Rel family members have been implicated as downstream targets of TCR and CD28 signalling events, particularly in the activation of the IL-2 promoter^{13,15,16}. The mechanism of NF- κ B activation after engagement of these cell-surface receptors has remained elusive. Phorbol esters induce NF- κ B in many different cell types and, in combination with Ca^{2+} ionophore, trigger the proliferation of T cells. In addition, phorbol ester-induced depletion of PKC abrogates TCR-mediated T-cell proliferation and NF- κ B activation¹⁷. On the basis of these results, it has been suggested that stimulation with TCR results in PKC-mediated activation of NF- κ B. Stimulation of T cells from *PKC- $\theta^{-/-}$* and wild-type mice with tumour-necrosis factor- α (TNF- α) or IL-1 resulted in strong activation of NF- κ B in nuclear extracts (Fig. 5b). The specificity of the electrophoretic mobility shift was confirmed by supershifting with antibodies against p50 or p65 (Fig. 5, and data not shown). In contrast, on crosslinking with anti-CD3 or anti-CD3 plus anti-CD28, there was a striking defect in NF- κ B activation in *PKC- $\theta^{-/-}$* T cells (Fig. 5a, b, and data not shown). T cells from the mutant mice also failed to activate NF- κ B in response to treatment with TPA. Although stimulation with TPA plus ionomycin did not restore a full proliferative response in mutant T cells, this treatment resulted in the activation of NF- κ B in the absence of PKC- θ (Fig. 5a). This finding suggests that recruitment of conventional PKCs (responsive to both diacylglycerol and cytoplasmic Ca^{2+}) can compensate for the absence of PKC- θ to restore at least part of the biological response downstream of the TCR/CD28 signal.

TCR-mediated signals required for selection events in the thymus and for the activation of peripheral T cells are thought to be fundamentally different, as they have different outcomes¹⁸. For example, antigens that stimulate the production of cytokines and the proliferation of mature T cells induce thymocyte apoptosis, whereas those that induce positive selection fail to activate in the periphery. Because T-cell development seemed unaffected in the absence of PKC- θ , we compared NF- κ B activation in thymocytes and splenic T cells. Surprisingly, activation of NF- κ B was found to be normal in *PKC- $\theta^{-/-}$* thymocytes treated with anti-CD3 or TPA

(Fig. 5c). This result suggests that a PKC isoenzyme distinct from PKC- θ is involved in TCR-initiated signalling events in thymocytes.

Before activation, NF- κ B subunits are sequestered in the cytosol by the negative regulator I κ B. To determine whether signalling to I κ B is defective in *PKC- $\theta^{-/-}$* T cells, we examined its degradation in mature T cells and thymocytes stimulated with anti-CD3 mAb or TNF- α . Consistent with the bandshift analysis, on CD3 crosslinking, I κ B- α was degraded in wild-type T cells but not in mature T cells lacking PKC- θ (Fig. 5d). As predicted, CD3-induced I κ B- α degradation was normal in thymocytes from wild-type and mutant mice, and no defect was observed in response to TNF- α in any of the cell types examined.

PKC- θ is a member of the 'novel' subclass of PKC family members, which also includes the δ , ϵ and η isoenzymes¹⁹. These kinases are activated by diacylglycerol, whereas the conventional subclass of PKCs, which include the α , β 1, β 2 and γ isoenzymes, require both diacylglycerol and Ca^{2+} for activation. The generation of diacylglycerol in response to a variety of signals results in plasma membrane recruitment and activation of PKCs. In addition, specific serine and threonine phosphorylation of conventional PKCs is required for them to acquire competence to recruitment and activation²⁰. It is not yet known whether PKC- θ activation is also dependent on similar phosphorylation events. However, recruitment of PKC- θ , but not other PKC isoenzymes, to the site of TCR interaction with major histocompatibility complex/peptide on antigen-presenting cells suggests that the local generation of diacylglycerol is not sufficient for membrane recruitment of PKC in T cells⁴. Specific signals are likely to engage and activate PKC- θ differentially after stimulation with anti-TCR, and the architecture of preformed signalling complexes probably accounts for the selective ability of this PKC isoenzyme to activate NF- κ B, AP-1

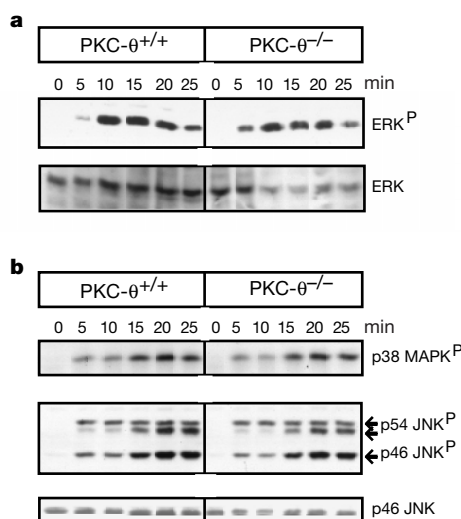


Figure 3 Normal MAP kinase activation in *PKC- $\theta^{-/-}$* T cells. **a**, Phospho-ERK immunoblot of T cells stimulated for the indicated durations with plate-bound anti-CD3E. As a control for loading, the blot was stripped and re-probed with anti-ERK antiserum. **b**, Activation of p38 MAPK and JNK in purified T cells preactivated with anti-CD3 and then rested and restimulated with a combination of anti-CD3 and anti-CD28. The same blot was probed with anti-phospho-p38 MAPK, anti-phospho-JNK and anti-JNK antibodies.

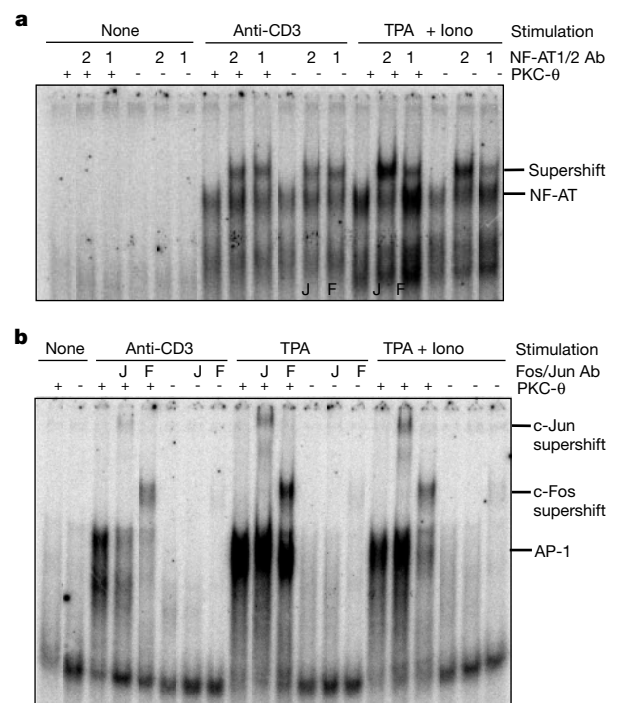
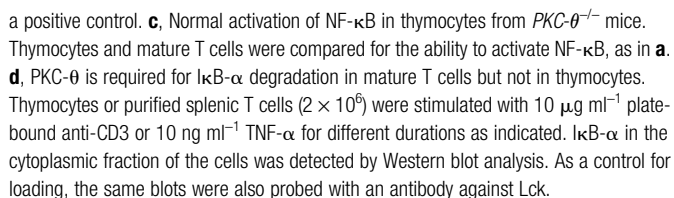


Figure 4 EMSA analysis of NF-AT and AP-1 in *PKC- $\theta^{-/-}$* T cells. Purified T cells were stimulated for 8 h with $10 \mu\text{g ml}^{-1}$ plate-bound anti-CD3 or 50 ng ml^{-1} TPA and 200 ng ml^{-1} ionomycin (Iono), and nuclear extracts were subjected to bandshift analysis. **a**, Normal translocation of NF-AT in activated *PKC- $\theta^{-/-}$* T cells. Antibodies (Ab) specific for NF-AT1 and NF-AT2 were employed in supershift assays as indicated. **b**, Defective activation of AP-1 in T cells lacking PKC- θ . Antibodies specific for c-Jun and c-Fos were used for supershift analysis.



selection of colonies with G418 and gancyclovir, homologous recombination events were confirmed by Southern blot analysis of DNA digested with *Kpn*I and *Bam*HI. Positive clones were used to generate chimaeric founder mice by microinjection into C57BL/6J blastocysts.

Antibodies against murine CD3e (145-2C11), CD28 (37.51), CD69 and CD25 and anti-human JNK were obtained from PharMingen. Antibodies against phospho-JNK and p38 MAPK were purchased from NEB, and anti-ERK1/2 was purchased from Promega. Antibodies specific for phospho-ERK (E-4), NF- κ B p50 (C-19), NF-AT1 (7A6), NF-AT2 (G1-D10), c-Jun (H-79), c-Fos (K-25) and I κ B- α (C-21) were obtained from Santa Cruz Biotechnology. An Lck-specific mAb was purchased from Transduction Labs and the T3.70 mAb was a gift from Hung-Sia Teh. TNF- α , IL-1 and IL-2 were obtained from Roche.

T cells isolated from lymph node or spleen were enriched to >80% purity with mouse T-cell enrichment columns (R&D Systems). For anti-CD3 stimulations, T cells (5×10^6) in 160 μ l proliferation medium (RPMI supplemented with 10% FCS, 50 μ M β -mercaptoethanol, 2 mM L-glutamine and 50 units ml^{-1} penicillin/streptomycin) were added in triplicate to 96-well plates precoated with anti-CD3 ϵ antibody (either 1.0 or 10.0 $\mu\text{g ml}^{-1}$). Where indicated, IL-2 (final concentration 20 units ml^{-1}) or anti-CD28 (1 $\mu\text{g ml}^{-1}$) were also added. TPA (10 ng ml^{-1}) was added either with ionomycin (125 ng ml^{-1}) or with soluble anti-CD28 (1 $\mu\text{g ml}^{-1}$). The T cells were incubated for 48 h and 100 μ l of supernatant was removed for IL-2 assays, before pulsing with 1 μCi of [^3H]thymidine for 16–20 h. For the MLR assays, T cells were purified from littermate mice (B6/129 background) and mixed in triplicate at various densities with mitomycin C-treated splenocytes from BABL/c mice, or with syngeneic splenocytes prepared in an identical manner. After 4 d of growth, the cells were pulsed for 12 h with 1 μCi [^3H]thymidine. Specific proliferation was calculated as the mean radioactive count in triplicate wells after subtraction of the baseline proliferation of syngeneic cultures.

IL-2 produced from the cultures was measured by quantifying the proliferation of the IL-2-dependent cell line HT-2. A standard curve was established by using recombinant IL-2.

Purified T cells were activated *ex vivo* on 96-well plates coated with anti-CD3 ($10 \mu\text{g ml}^{-1}$). For the assay shown in Fig. 3b, T cells were first incubated in 10% FCS in RPMI on anti-

A 129Sv/J mouse genomic clone containing the exon encoding the ATP-binding site of the kinase domain, which corresponds to exon 11 in the human *PKC-θ* gene, was used to prepare the targeting vector. A 1.2-kilobase (kb) genomic fragment located 5' to the exon was amplified by the polymerase chain reaction and cloned between the *Clal* and *XbaI* sites in the PL2-Neo plasmid polylinker (a gift from H. Gu and K. Rajewsky). At the other end of the Neo selection marker, a 6-kb 3' *Bam*HI fragment, with a flanking HSV-TK gene, was inserted. E14 ES cells were transfected with linearized targeting vector and, after the

CD3-coated tissue culture dishes. After 16 h, cells were gently detached, rested for 3–5 h on normal tissue-culture dishes, and then activated for various durations on 96-well plates containing co-immobilized anti-CD3 and anti-CD28 (5 $\mu\text{g ml}^{-1}$ each). Proteins were extracted by the addition of an equal volume of 2 $\times$ reducing sample buffer and, after being boiled, (0.5–1.0) $\times 10^6$ cell equivalents were added to the lanes of 10% polyacrylamide gels. Proteins were transferred to Immobilon-P (Millipore) and probed with phospho-specific MAP kinase antibodies for detection by chemiluminescence (Pierce).

Nuclear extract preparation and electrophoretic mobility-shift analysis

Purified T cells (10^6) were washed in PBS and resuspended in 1 ml of 10 mM HEPES pH 7.9, 10 mM KCl, 0.1 mM EDTA and Complete protease inhibitors (Roche). The cells were incubated on ice for 15 min. Nonidet P40 was added to a final concentration of 0.5%, the cells were vortex-mixed vigorously, and the mixture was centrifuged for 30 s. The nuclear pellets were resuspended in 40 μl of 20 mM HEPES pH 7.9, 0.4 M NaCl, 1 mM EDTA and protease inhibitors, and the tube was rocked for 15 min at 4 °C. After centrifugation for 5 min, the supernatant was collected and glycerol was added to a final concentration of 50%; 2 μl of each extract was subjected to bandshift analysis. Electrophoretic mobility-shift analyses (EMSAs) were performed with the following oligonucleotides and binding buffers: NFAT oligonucleotide: 5'-GCCCAAAGAGGAAAA TTTGTTTCATACAG-3'; NFAT 5 $\times$ binding buffer: 50 mM Tris-HCl pH 7.5, 500 mM KCl, 2.5 mM MgCl₂, 0.5 mM EDTA, 50% glycerol, 250 $\mu\text{g ml}^{-1}$ poly (dI:dC), 5 mM dithiothreitol and 1 mg ml⁻¹ BSA; NF- κ B oligonucleotide: 5'-ACCAAGAGGGATTTCACCT AAATC-3'; NF- κ B 5 $\times$ binding buffer: 25 mM Tris-HCl pH 7.5, 5 mM EDTA, 250 $\mu\text{g ml}^{-1}$ poly (dI:dC), 5 mM dithiothreitol and 1 mg ml⁻¹ BSA; AP-1 oligonucleotide: 5'-CGCTT GATGACTCAGCCGAA-3'; AP-1 5 $\times$ binding buffer: same as that for NFAT. 5 $\times 10^5$ c.p.m. of labelled probe was used in each reaction, and bandshifts were resolved on 5% polyacrylamide gels in 0.5 $\times$ TBE running buffer.

Immunizations

KLH (20 $\mu\text{g ml}^{-1}$) was allowed to adsorb on 1 mg of aluminium hydroxide for 1 h at room temperature. After washing unadsorbed KLH with PBS, the KLH-aluminium hydroxide mixture was resuspended in 200 μl of PBS and injected intraperitoneally. Splenocytes recovered 2 weeks later were stimulated at 2×10^5 cells per well (in triplicate) with 0, 0.1, 1 or 10 $\mu\text{g ml}^{-1}$ KLH for 3 d. Cells were then pulsed with 1 μCi of [³H]thymidine for 1 d.

Received 4 January; accepted 8 February 2000.

- Monks, C. R., Freiberg, B. A., Kupfer, H., Sciaky, N. & Kupfer, A. Three-dimensional segregation of supramolecular activation clusters in T cells. *Nature* **395**, 82–86 (1998).
- Grakoui, A. *et al.* The immunological synapse: a molecular machine controlling T cell activation. *Science* **285**, 221–227 (1999).
- Valge, V. E., Wong, J. G., Datlof, B. M., Sinskey, A. J. & Rao, A. Protein kinase C is required for responses to T cell receptor ligands but not to interleukin-2 in T cells. *Cell* **55**, 101–112 (1988).
- Monks, C. R., Kupfer, H., Tamir, I., Barlow, A. & Kupfer, A. Selective modulation of protein kinase C- θ during T-cell activation. *Nature* **385**, 83–86 (1997).
- Werlen, G., Jacinto, E., Xia, Y. & Karin, M. Calcineurin preferentially synergizes with PKC- θ to activate JNK and IL-2 promoter in T lymphocytes. *EMBO J.* **17**, 3101–3111 (1998).
- Ghaffari-Tabrizi, N. *et al.* Protein kinase C θ , a selective upstream regulator of JNK/SAPK and IL-2 promoter activation in Jurkat T cells. *Eur. J. Immunol.* **29**, 132–142 (1999).
- Kisielow, P., Bluthmann, H., Staerz, U. D., Steinmetz, M. & von Boehmer, H. Tolerance in T-cell-receptor transgenic mice involves deletion of nonmature CD4⁺8⁺ thymocytes. *Nature* **333**, 742–746 (1988).
- Meller, N., Altman, A. & Isakov, N. New perspectives on PKC θ , a member of the novel subfamily of protein kinase C. *Stem Cells* **16**, 178–192 (1998).
- Smith, K. A. The interleukin 2 receptor. *Annu. Rev. Cell Biol.* **5**, 397–425 (1989).
- Genot, E. M., Parker, P. J. & Cantrell, D. A. Analysis of the role of protein kinase C- α , - ϵ , and - ζ in T cell activation. *J. Biol. Chem.* **270**, 9833–9839 (1995).
- Weiss, A. & Littman, D. R. Signal transduction by lymphocyte antigen receptors. *Cell* **76**, 263–274 (1994).
- Weiss, L. *et al.* Regulation of c-Jun NH(2)-terminal kinase (jnk) gene expression during T cell activation. *J. Exp. Med.* **191**, 139–146 (2000).
- Shapiro, V. S., Truitt, K. E., Imboden, J. B. & Weiss, A. CD28 mediates transcriptional upregulation of the interleukin-2 (IL-2) promoter through a composite element containing the CD28RE and NF-IL-2B AP-1 sites. *Mol. Cell. Biol.* **17**, 4051–4058 (1997).
- Northrop, J. P. *et al.* NF-AT components define a family of transcription factors targeted in T-cell activation. *Nature* **369**, 497–502 (1994).
- Jain, J., Loh, C. & Rao, A. Transcriptional regulation of the IL-2 gene. *Curr. Opin. Immunol.* **7**, 333–342 (1995).
- Jamieson, C., McCaffrey, P. G., Rao, A. & Sen, R. Physiologic activation of T cells via the T cell receptor induces NF- κ B. *J. Immunol.* **147**, 416–420 (1991).
- Jain, J., Valge-Archer, V. E., Sinskey, A. J. & Rao, A. The AP-1 site at -150 bp, but not the NF- κ B site, is likely to represent the major target of protein kinase C in the interleukin 2 promoter. *J. Exp. Med.* **175**, 853–862 (1992).
- Sloan-Lancaster, J. & Allen, P. M. Altered peptide ligand-induced partial T cell activation: molecular mechanisms and role in T cell biology. *Annu. Rev. Immunol.* **14**, 1–27 (1996).
- Newton, A. C. Regulation of protein kinase C. *Curr. Opin. Cell Biol.* **9**, 161–167 (1997).
- Edwards, A. S., Faux, M. C., Scott, J. D. & Newton, A. C. Carboxyl-terminal phosphorylation regulates the function and subcellular localization of protein kinase C β II. *J. Biol. Chem.* **274**, 6461–6468 (1999).
- Kong, Y. Y. *et al.* Vav regulates peptide-specific apoptosis in thymocytes. *J. Exp. Med.* **188**, 2099–2111 (1998).
- Fischer, K. D. *et al.* Vav is a regulator of cytoskeletal reorganization mediated by the T-cell receptor. *Curr. Biol.* **8**, 554–562 (1998).

- Holsinger, L. J. *et al.* Defects in actin-cap formation in Vav-deficient mice implicate an actin requirement for lymphocyte signal transduction. *Curr. Biol.* **8**, 563–572 (1998).
- Kawakami, Y. *et al.* Activation and interaction with protein kinase C of a cytoplasmic tyrosine kinase, Itk/Tsk/Emt, on Fc ϵ R1 cross-linking on mast cells. *J. Immunol.* **155**, 3556–3562 (1995).
- Zandi, E. & Karin, M. Bridging the gap: composition, regulation, and physiological function of the I κ B kinase complex. *Mol. Cell. Biol.* **19**, 4547–4551 (1999).
- Esslinger, C. W., Jongeneel, C. V. & MacDonald, H. R. Survival-independent function of NF- κ B/Rel during late stages of thymocyte differentiation. *Mol. Immunol.* **35**, 847–852 (1998).
- Boothby, M. R., Mora, A. L., Scherer, D. C., Brockman, J. A. & Ballard, D. W. Perturbation of the T lymphocyte lineage in transgenic mice expressing a constitutive repressor of nuclear factor (NF)- κ B. *J. Exp. Med.* **185**, 1897–1907 (1997).
- Hettmann, T., DiDonato, J., Karin, M. & Leiden, J. M. An essential role for nuclear factor κ B in promoting double positive thymocyte apoptosis. *J. Exp. Med.* **189**, 145–158 (1999).

Acknowledgements

We thank S. Marmon for preparing fusion proteins used to generate antisera; M. Brown for analysis of the antisera; B. Alfonso and W. O'Brien for genotyping; D. Alpert, M. Rincon and R. Davis for advice on the JNK assays; and E. Skolnik, A. Zychlinsky, Y. Zou and D. Nuttmaz for their critical reading of the manuscript. C.W.A. is supported by a fellowship award from the Medical Research Council of Canada. Z.S. is supported by an NIH Molecular Oncology and Immunology training grant. E.M.S. is a Howard Hughes Medical Institute–NIH Research Scholar. P.L.S. is supported in part by a Searle Scholar award, and A.K. is supported by a grant from the NIH. W.E. is an Associate and D.R.L. is an Investigator of the Howard Hughes Medical Institute.

Correspondence and requests for materials should be addressed to D.R.L. (e-mail: littman@saturn.med.nyu.edu).

Control of T_H2 polarization by the chemokine monocyte chemoattractant protein-1

Long Gu, Susan Tseng, Renée M. Horner, Carmen Tam, Massimo Loda & Barrett J. Rollins

Department of Adult Oncology, Dana-Farber Cancer Institute, Brigham & Women's Hospital, and Harvard Medical School, 44 Binney Street, Boston, Massachusetts 02115, USA

Activated T lymphocytes differentiate into effector cells tailored to meet disparate challenges to host integrity¹. For example, type 1 and type 2 helper (T_H1 and T_H2) cells secrete cytokines that enhance cell-mediated and humoral immunity, respectively. The chemokine monocyte chemoattractant protein-1 (MCP-1) can stimulate interleukin-4 production² and its overexpression is associated with defects in cell-mediated immunity³, indicating that it might be involved in T_H2 polarization. Here we show that MCP-1-deficient mice are unable to mount T_H2 responses. Lymph node cells from immunized MCP-1^{-/-} mice synthesize extremely low levels of interleukin-4, interleukin-5 and interleukin-10, but normal amounts of interferon- γ and interleukin-2. Consequently, these mice do not accomplish the immunoglobulin subclass switch that is characteristic of T_H2 responses and are resistant to *Leishmania major*. These effects are direct rather than due to abnormal cell migration, because the trafficking of naive T cells is undisturbed in MCP-1^{-/-} mice despite the presence of MCP-1-expressing cells in secondary lymphoid organs of wild-type mice. Thus, MCP-1 influences both innate immunity, through effects on monocytes, and adaptive immunity, through control of T helper cell polarization.

MCP-1 is a CC chemokine that attracts monocytes, memory T lymphocytes and natural killer cells *in vitro*^{4–8}. Several models of transgenic expression in the mouse have validated it as a predominantly monocytic chemoattractant *in vivo*^{3,9–11}. Despite the existence of other CC chemokines that attract monocytes *in vitro* with

the same potency, MCP-1-deficient mice demonstrate that it is solely responsible for monocyte recruitment in several inflammatory settings¹². There is similar evidence for an essential pathogenic role of MCP-1 in disease models such as atherosclerosis^{13,14} and asthma¹⁵, in which MCP-1 is required for the accumulation of macrophages in diseased tissue. Because of these findings and the observation that inflammatory stimuli induce MCP-1 expression, MCP-1 has been considered to be an inflammatory chemokine that has little effect on adaptive immunity or immune-cell trafficking.

However, several lines of evidence indicate that MCP-1 might influence T-cell immunity. First, T lymphocytes express the MCP-1 receptor CCR2 (ref. 16), and MCP-1 is a potent chemoattractant for memory T cells *in vitro*^{6,16} (although T cells are nearly absent from transgenic MCP-1-induced infiltrates *in vivo*). Second, MCP-1 expression is associated with the development of polarized TH2 responses^{12,17}, and MCP-1 enhances the secretion of interleukin-4 (IL-4) by T cells^{2,18}. Third, in immune-mediated diseases with a TH2 character, such as asthma, MCP-1 is expressed at high levels and its neutralization in animal models ameliorates disease¹⁵. Last, other chemokines and their receptors are linked to specific responses of T-helper cells^{19,20}. These observations prompted our analysis of MCP-1's influence on the development of polarized immune responses *in vivo*.

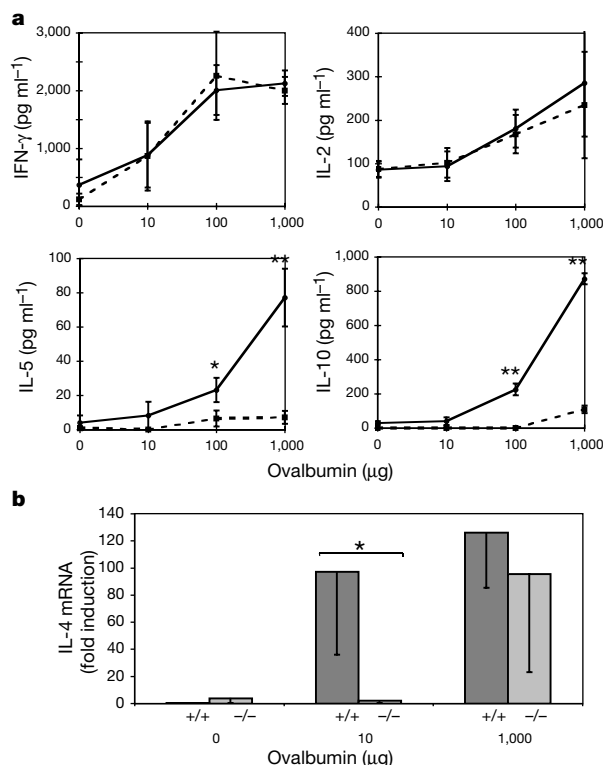


Figure 1 Effect of MCP-1 on cytokine synthesis. **a**, Seven days after immunization with TNP-Ova, single-cell suspensions were prepared from draining lymph nodes of *MCP-1*^{+/+} and *MCP-1*^{-/-} mice. Cells were challenged *in vitro* with increasing concentrations of ovalbumin, and concentrations of the indicated cytokines were determined by ELISA every 24 h for 3 d. Dose responses are shown for the day on which the highest cytokine levels were observed (day 2 for interferon-γ and IL-2, day 3 for IL-5 and IL-10). Values are means for 6 mice per group and are representative of two independently performed experiments. Solid line, *MCP-1*^{+/+}; dotted line, *MCP-1*^{-/-}. Asterisk, *P* < 0.05 compared with *MCP-1*^{-/-} mice; two asterisks, *P* < 0.005 compared with *MCP-1*^{-/-} mice. **b**, Cells from *MCP-1*^{+/+} (+/+) or *MCP-1*^{-/-} (-/-) mice were challenged *in vitro* with the indicated concentrations of ovalbumin for 3 d, and IL-4 mRNA content was determined by quantitative PCR and normalized to GAPDH mRNA levels. Induction is normalized to untreated *MCP-1*^{+/+} cells, and values represent the means of 3 mice per group. Asterisk, *P* = 0.05 for the difference between *MCP-1*^{+/+} and *MCP-1*^{-/-} mice.

To examine the potential for *MCP-1*^{-/-} mice to develop TH2 responses, we immunized mice with trinitrophenol-derivatized ovalbumin (TNP-Ova) in incomplete Freund's adjuvant. One week later, draining lymph nodes were isolated and single-cell suspensions were challenged with increasing concentrations of ovalbumin. Figure 1a shows that cells from *MCP-1*^{+/+} and *MCP-1*^{-/-} mice secreted the same amounts of interferon-γ and IL-2, suggesting that the TH1 component of this response was intact. Consistent with that observation was the finding that IL-12 levels were the same in supernatants from both cultures (~30 pg ml⁻¹). In contrast, cells from *MCP-1*^{-/-} mice secreted extremely low levels of IL-5 and IL-10, whereas cells from wild-type mice secreted substantial amounts.

Because lymph node cells from immunized wild-type C57Bl/6 mice secreted undetectable amounts of IL-4, we measured the expression of IL-4 messenger RNA by polymerase chain reaction with reverse transcription (RT-PCR) (Fig. 1b). Quantitative analysis showed that *MCP-1*^{-/-} mice expressed 50-fold less IL-4 messenger RNA than did wild-type mice in response to intermediate doses of ovalbumin (10 μg ml⁻¹). At higher ovalbumin challenge doses (1000 μg ml⁻¹), however, this deficiency could be overcome, which was not true of IL-5 and IL-10 (Fig. 1a), suggesting the possibility that MCP-1's effects on TH2 polarization may be partially independent of IL-4.

We examined the physiological consequences of this deficiency in cytokine secretion by testing whether or not *MCP-1*^{-/-} mice could accomplish immunoglobulin subclass switching characteristic of TH2 responses. First, however, Fig. 2a shows that the ratio of total serum IgG₁ to IgG_{2a} in naive *MCP-1*^{-/-} mice was already decreased 10-fold as compared with wild-type mice. The ratio of total IgG₁ to IgG_{2b} was decreased twofold, which, although a less substantial difference, was still highly significant. These altered ratios were due both to decreases in IgG₁ and to increases in IgG_{2a} and IgG_{2b} in *MCP-1*^{-/-} mice as compared with wild-type mice. Thus, in the

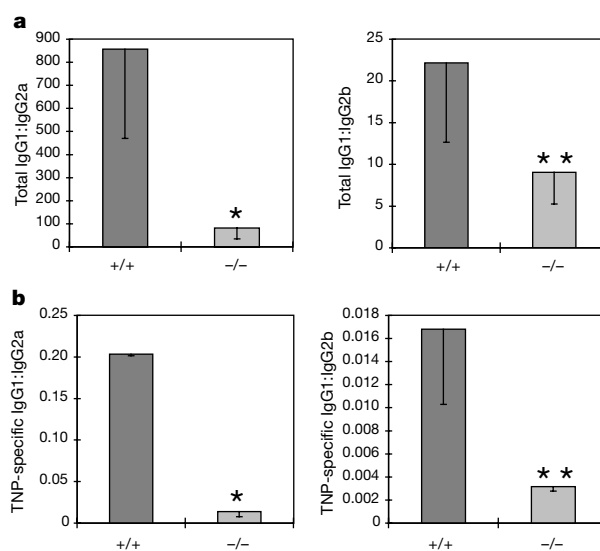


Figure 2 Effect of MCP-1 on IgG subclass concentrations. **a**, Serum was collected from *MCP-1*^{+/+} (+/+) and *MCP-1*^{-/-} (-/-) mice, and total IgG subclass concentrations were determined by ELISA. To normalize for variation between animals, the data are shown as ratios of IgG₁ to IgG_{2a} and IgG₁ to IgG_{2b}. Means were derived from three or four animals in each group; these data are typical of two independently performed experiments. Asterisk, *P* < 0.008 compared with *MCP-1*^{+/+} mice; two asterisks, *P* < 0.05 compared with *MCP-1*^{+/+} mice. **b**, TNP-specific IgG subclass ratios were determined in mice 7 d after immunization with TNP-Ova. Means were derived from three or four animals in each group; these data are typical of two independently performed experiments. Asterisk, *P* < 2 × 10⁻⁷ compared with *MCP-1*^{+/+} mice; two asterisks, *P* < 0.008 compared with *MCP-1*^{+/+} mice.

absence of MCP-1, mice have total immunoglobulin subclass concentrations indicative of a baseline T_H1 bias. After immunization, the T_H1 -skewed subclass ratios observed in naive mice persisted in TNP-specific immunoglobulin (Fig. 2b). The ratio of TNP-specific IgG₁ to IgG_{2a} was decreased 10-fold and the ratio of TNP-specific IgG₁ to IgG_{2b} was decreased 4-fold in $MCP-1^{-/-}$ mice in comparison with wild-type mice. Because IL-4 stimulates the IgG₁ switch and inhibits the IgG_{2a} and IgG_{2b} switch, the fact that high concentrations of antigen could overcome the IL-4 mRNA blockade *in vitro* (Fig. 2b) seems to be physiologically irrelevant in this setting.

To support further the notion that $MCP-1^{-/-}$ mice are T_H2 deficient, we tested their susceptibility to infection by *Leishmania major*. Because Balb/c mice mount a T_H2 response to this parasite, they succumb to infection. By 7 weeks after the injection of 2×10^6 promastigotes in the hind footpad, $MCP-1^{-/-}$ Balb/c mice had significantly less footpad swelling than wild-type Balb/c mice (Fig. 3) and their lesions were stabilizing, whereas those in wild-type mice continued to progress to the point that it was necessary to kill many of the wild-type controls by 9–10 weeks. However, lesions in $MCP-1^{-/-}$ mice had not yet healed by this time, suggesting that their phenotype was intermediate between completely susceptible and completely resistant strains.

Although the development of a T_H2 response could be due to a direct effect of MCP-1 on T cells, it is also possible that MCP-1 attracts precursor cells to a microenvironment containing a mixture of cells and cytokines that are actually responsible for T_H2 polarization. In that case, naive precursor cells might require MCP-1 if they are to travel to T-cell zones in secondary lymphoid organs where they encounter antigen. To test this idea, we isolated naive T lymphocytes from a congenic mouse strain carrying the Ly-5.1 (CD45.1) allele and injected them intravenously into wild-type C57Bl/6 or $MCP-1^{-/-}$ mice in a C57Bl/6 background, which carry the Ly-5.2 (CD45.2) allele. Five hours later, exogenous T cells migrating to the T-cell zones of spleens and lymph nodes were identified by using anti-Ly-5.1 (Fig. 4a). Image analysis showed that the same numbers of exogenous T cells had accumulated in the T-cell zones of wild-type and $MCP-1^{-/-}$ mice (Fig. 4b), indicating the absence of a T-cell trafficking defect in $MCP-1^{-/-}$ mice.

Nevertheless, *in situ* hybridization revealed the presence of cells in secondary lymphoid organs that constitutively express MCP-1. Figure 5 shows that scattered cells in the splenic periarteriolar lymphatic sheaths (PALS) expressed MCP-1. In lymph nodes, a similarly scattered distribution of MCP-1-expressing cells was found in the medulla (Fig. 5f). Seven days after immunization, there was a general increase in low-level MCP-1 expression throughout the PALS, and cells expressing higher levels of MCP-1 extended

farther from the arteriole (Fig. 5d). Although we have not yet identified the cell types responsible for this expression, they are clearly localized to the T-cell zone and do not have the morphology of macrophages or dendritic cells. There are therefore MCP-1-expressing cells within secondary lymphoid organs that are in locations appropriate for taking a role in the direct stimulation of polarized T-cell responses.

Despite MCP-1's effects on monocytes and macrophages, which might have implied a role in T_H1 -polarized responses, our data indicate instead that it is required for the development of T_H2 responses. The resistance of $MCP-1^{-/-}$ mice to *Mycobacterium tuberculosis*¹² is consistent with their fully functional T_H1 responses. The severe T_H2 defect in $MCP-1^{-/-}$ mice is unlikely to be due to adventitious compensatory changes in response to MCP-1 disruption because animals made acutely MCP-1-deficient by using neutralizing antibodies also had difficulty in mounting full T_H2 responses¹⁷. Nevertheless, this deficiency contrasts with a reported T_H1 defect in $CCR2^{-/-}$ mice, which lack the only receptor for MCP-1 that has been cloned so far^{21,22}. This discrepancy might not be surprising given the fact that CCR2 has at least two additional high-affinity ligands in the mouse (MCP-3 and MCP-5), which, in an appropriate context, might stimulate T_H1 polarization. Furthermore, there is biological evidence for another MCP-1 receptor²³, and the receptor known as D6 binds MCP-1, although it does not signal²⁴. MCP-1 might exert its effects on T_H2 polarization through these receptors rather than CCR2. Lastly, the chemokine receptor CCR4 has been reported to bind MCP-1 with high affinity²⁵.

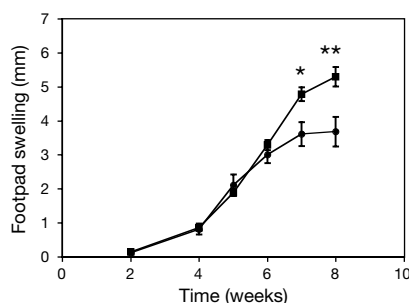


Figure 3 Resistance of $MCP-1^{-/-}$ mice to *Leishmania major*. Promastigotes (2×10^6) were injected into the right hind footpads of six wild-type Balb/c mice and eight $MCP-1^{-/-}$ mice in a Balb/c background. Right and left footpad thicknesses were measured weekly and footpad swelling was determined as the difference between right and left footpad thicknesses. Circles, wild-type; squares, $MCP-1^{-/-}$; bars indicate s.e.m. Asterisk, $P = 0.05$ compared with wild-type mice; two asterisks, $P = 0.007$ compared with wild-type mice.

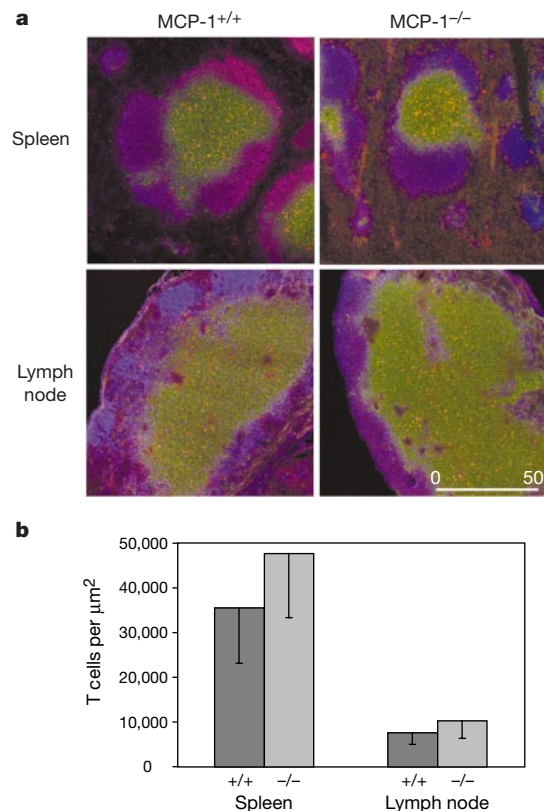


Figure 4 Effect of MCP-1 on T-lymphocyte trafficking. **a**, Purified T lymphocytes from wild-type CD45.1⁺ (Ly-5.1⁺) mice were injected intravenously into CD45.2⁺ (Ly-5.2⁺) $MCP-1^{+/+}$ and $MCP-1^{-/-}$ mice. Five hours later, spleens and lymph nodes were isolated and processed for immunohistochemistry. B-cell areas are stained purple, T-cell areas are stained green, and injected T cells are stained red. Scale bar, 50 μm. **b**, The area of the T-cell zone and the number of injected T cells per μm² of T-cell zone were determined by image analysis of confocal microscopy images. Means and s.e.m. are shown for three wild-type and three $MCP-1^{-/-}$ mice.

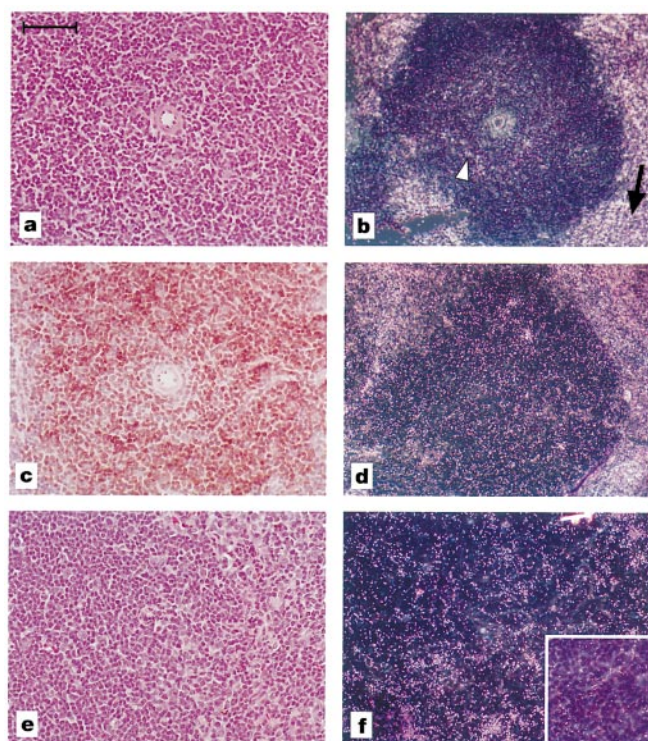


Figure 5 Expression of MCP-1 in spleen and lymph nodes. Spleens and lymph nodes from naive C57Bl/6 mice and C57Bl/6 mice 7 d after immunization were analysed for MCP-1 expression by *in situ* hybridization. **a**, Spleen from naive mouse stained with haematoxylin and eosin; **b**, spleen from naive mouse hybridized with antisense MCP-1 probe; **c**, spleen from naive mouse stained with anti-CD3; **d**, spleen from immunized mouse hybridized with antisense MCP-1 probe; **e**, lymph node from naive mouse stained with haematoxylin and eosin; **f**, lymph node from naive mouse hybridized with antisense MCP-1 probe (inset,

serial section hybridized with sense MCP-1 probe). The white arrowhead in **b** indicates specific probe hybridization; the black arrow indicates areas of birefringence in the red pulp that did not show specific hybridization under bright-field examination. Panels **a**, **b**, **c**, **e** and **f** are serial sections. Hybridization of tissues from *MCP-1*^{-/-} mice with the antisense probe showed no specific hybridization, indicating that the murine MCP-1 probe did not cross-hybridize with other chemokine mRNAs. Scale bar, 100 μ m.

Although others have not reproduced this finding²⁶, CCR4 is preferentially expressed by T_H2 cells^{19,20}, and we cannot exclude the possibility that MCP-1 might produce some of its effects on T-cell polarization by using this receptor.

MCP-1's influence on memory/effector T-cell heterogeneity has important implications for pathogenesis. For example, MCP-1 is thought to have a pathobiological role in inflammatory disorders because of its ability to attract monocytes. However, neutralization of MCP-1 in rodent models of asthma, for example, accomplishes much more than simply preventing the appearance of macrophages in the air spaces¹⁵. It also reduces the total cellularity of the bronchoalveolar lavage fluid and lessens airway hyperreactivity. If MCP-1's sole activity were monocyte chemoattraction, its neutralization would be unlikely to have such wide-ranging effects. But, given its influence on T_H2 polarization, MCP-1 could affect both proximal (that is, immune T_H2-driven) and distal (that is, macrophage effector) steps in asthma pathogenesis. The same dual control of inflammatory pathology might occur in human asthma and in other diseases with high levels of MCP-1 expression, such as atherosclerosis or dementia associated with human immunodeficiency virus^{13,27,28}. The present work shows that a single chemokine can affect two disparate and essential pathways in disease, making MCP-1 a very attractive target for anti-inflammatory drugs. □

Methods

Immunization

TNP-Ova was prepared by incubating a 5:1 mixture of 1% ovalbumin (Sigma) and 5% trinitrobenzenesulphonic acid (Sigma) at room temperature for 4 h, followed by extensive dialysis against PBS buffer. A total of 200 μ g of TNP-Ova emulsified in incomplete Freund's adjuvant was injected subcutaneously in three to five sites on the back and in both hind footpads of wild-type C57Bl/6 mice or *MCP-1*^{-/-} mice in a C57Bl/6 background.

Cytokine measurements

Single-cell suspensions were prepared from lymph nodes 7 d after immunization with TNP-Ova. Cells were added to multiwell tissue-culture plates at 2×10^6 cells ml⁻¹ in RPMI 1640 supplemented with 10% fetal bovine serum. Ovalbumin was added to each well at the indicated concentration and supernatants were recovered every 24 h for 3 d. Cytokine concentrations were determined by enzyme-linked immunosorbent assay (ELISA) (R&D Systems). For RT-PCR analysis, RNA was isolated by using RNeasy (Qiagen). Total RNA (2 μ g) was reverse-transcribed and the cDNA quantified by real-time PCR with the TaqMan system (PE Corporation). Primer and probe sequences for IL-4 and glyceraldehyde-3-phosphate dehydrogenase (GAPDH) are available on request. Values for IL-4 mRNA were normalized to the relative quantity of GAPDH mRNA in each sample.

IgG subclass measurements

To measure total IgG subclass concentrations, serum from naive or immunized mice was incubated in plates precoated with capturing antibodies specific for each subclass (PharMingen) for 2 h at room temperature. After five washes with PBS, 0.05% Tween 20, the appropriately matched biotinylated anti-subclass antibody was added to each well and incubated for 2 h at room temperature. After another five washes with PBS, 0.05% Tween 20, streptavidin-conjugated horseradish peroxidase was added, the plates were developed with 3-ethylbenzthiazoline-6-sulphonic acid and absorbance was read at 405 nm. Measurement of TNP-specific IgG subclass concentrations was performed in a similar manner except that the ELISA plates were precoated with TNP-BSA rather than capture antibodies.

Leishmania major infection

Promastigotes were provided by A. Satoskar and J. David (Harvard School of Public Health). *L. major* parasites (2×10^6) were injected into the hind footpads of 4–6-week-old male wild-type Balb/c mice or *MCP-1*^{-/-} mice in a Balb/c background, and footpad thickness was measured weekly with a Peacock dial gauge.

T-cell trafficking.

T lymphocytes were isolated from lymph nodes and spleens of 8–12-week-old male C57Bl/6 mice congenic for CD45.1 (a gift from L. Clayton and E. Reinherz, Dana-Farber Cancer Institute) with a negative purification method (StemCell Technologies). Cell preparations (>95% T cells) were adjusted to 2×10^7 ml⁻¹ and injected into the tail veins of

8–12-week-old male wild-type or *MCP-1^{-/-}* mice. Five hours after injection, spleens and lymph nodes were embedded in Tissue-Tek optimum cutting temperature compound (VWR) sectioned by cryostat and fixed in acetone. Sections were incubated for 1 h with Cychrome-conjugated rat anti-mouse CD45R/B220, fluorescein isothiocyanate-conjugated rat anti-mouse CD90.2 (Thy-1.2) and phycoerythrin-conjugated mouse anti-mouse CD45.1 (Ly-5.1), each diluted 1:100 in PBS (PharMingen). After three washes with PBS, 0.05% Tween 20, slides were mounted and examined by confocal microscopy. Image analysis was performed with Image Pro software (Media Cybernetics).

Immunohistochemistry and *in situ* hybridization

Spleens and lymph nodes were fixed in 4% paraformaldehyde in PBS and embedded in paraffin. Sections were processed for immunohistochemistry by using rat anti-human CD3, which crossreacts with murine CD3 (Harlan Bioproducts). A rat myeloma protein of identical isotype was used as a primary antibody control. Slides were incubated with biotinylated secondary antibodies and avidin–biotin–complexed horseradish peroxidase (Vector Laboratories), and developed with diaminobenzamide. For *in situ* hybridization, probes were generated by *in vitro* transcription of pcJE-1 (ref. 29) in the presence of [³⁵S]UTP. Transcription, hybridization and development were performed as described³⁰.

Received 20 December 1999; accepted 26 January 2000.

1. Abbas, A. K., Murphy, K. M. & Sher, A. Functional diversity of helper T lymphocytes. *Nature* **383**, 787–793 (1996).
2. Karpus, W. J. *et al.* Differential CC chemokine-induced enhancement of T helper cell cytokine production. *J. Immunol.* **158**, 4129–4136 (1997).
3. Rutledge, B. J. *et al.* High level monocyte chemoattractant protein-1 expression in transgenic mice increases their susceptibility to intracellular pathogens. *J. Immunol.* **155**, 4838–4843 (1995).
4. Matsushima, K., Larsen, C. G., DuBois, G. C. & Oppenheim, J. J. Purification and characterization of a novel monocyte chemotactic and activating factor produced by a human myelomonocytic cell line. *J. Exp. Med.* **169**, 1485–1490 (1989).
5. Yoshimura, T. *et al.* Purification and amino acid analysis of two human glioma-derived monocyte chemoattractants. *J. Exp. Med.* **169**, 1449–1459 (1989).
6. Carr, M. W., Roth, S. J., Luther, E., Rose, S. S. & Springer, T. A. Monocyte chemoattractant protein 1 acts as a T-lymphocyte chemoattractant. *Proc. Natl. Acad. Sci. USA* **91**, 3652–3656 (1994).
7. Allavena, P. *et al.* Induction of natural killer cell migration by monocyte chemotactic protein-1, -2 and -3. *Eur. J. Immunol.* **24**, 3233–3236 (1994).
8. Maghazachi, A. A., al-Aoukaty, A. & Schall, T. J. C-C chemokines induce the chemotaxis of NK and IL-2-activated NK cells. Role for G proteins. *J. Immunol.* **153**, 4969–4977 (1994).
9. Fuentes, M. E. *et al.* Controlled recruitment of monocytes/macrophages to specific organs through transgenic expression of MCP-1. *J. Immunol.* **155**, 5769–5776 (1995).
10. Nakamura, K., Williams, I. R. & Kupper, T. S. Keratinocyte-derived monocyte chemoattractant protein 1 (MCP-1): analysis in a transgenic model demonstrates MCP-1 can recruit dendritic and Langerhans cells to skin. *J. Invest. Dermatol.* **105**, 635–643 (1995).
11. Gunn, M. D., Nelken, N. A., Liao, X. & Williams, L. T. Monocyte chemoattractant protein-1 is sufficient for the chemotaxis of monocytes and lymphocytes in transgenic mice but requires an additional stimulus for inflammatory activation. *J. Immunol.* **158**, 376–383 (1997).
12. Lu, B. *et al.* Abnormalities in monocyte recruitment and cytokine expression in monocyte chemoattractant protein 1-deficient mice. *J. Exp. Med.* **187**, 601–608 (1998).
13. Gu, L. *et al.* Absence of monocyte chemoattractant protein-1 reduces atherosclerosis in low density lipoprotein receptor-deficient mice. *Mol. Cell* **2**, 275–281 (1998).
14. Gosling, J. *et al.* MCP-1 deficiency reduces susceptibility to atherosclerosis in mice that overexpress human apolipoprotein B: a critical role for chemokines in monocyte recruitment and foam cell formation. *J. Clin. Invest.* **103**, 773–778 (1999).
15. Gonzalo, J. A. *et al.* The coordinated action of CC chemokines in the lung orchestrates allergic inflammation and airway hyperresponsiveness. *J. Exp. Med.* **188**, 157–167 (1998).
16. Loetscher, P., Seitz, M., Baggiolini, M. & Moser, B. Interleukin-2 regulates CC chemokine receptor expression and chemotactic responsiveness in T lymphocytes. *J. Exp. Med.* **184**, 569–577 (1996).
17. Chensue, S. W. *et al.* Monocyte chemotactic protein expression during schistosoma egg granuloma formation. Sequence of production, localization, contribution, and regulation. *Am. J. Pathol.* **146**, 130–138 (1995).
18. Karpus, W. J., Kennedy, K. J., Kunkel, S. L. & Lukacs, N. W. Monocyte chemotactic protein 1 regulates oral tolerance induction by inhibition of T helper cell 1-related cytokines. *J. Exp. Med.* **187**, 733–741 (1998).
19. Sallusto, F., Lanzavecchia, A. & Mackay, C. R. Chemokines and chemokine receptors in T-cell priming and Th1/Th2-mediated responses. *Immunol. Today* **19**, 568–574 (1998).
20. Bonecchi, R. *et al.* Differential expression of chemokine receptors and chemotactic responsiveness of type 1 T helper cells (Th1s) and Th2s. *J. Exp. Med.* **187**, 129–134 (1998).
21. Boring, L. *et al.* Impaired monocyte migration and reduced type 1 (Th1) cytokine responses in C-C chemokine receptor 2 knockout mice. *J. Clin. Invest.* **100**, 2552–2561 (1997).
22. Warmington, K. S. Effect of C-C chemokine receptor 2 (CCR2) knockout on type-2 (schistosomal antigen-elicited) pulmonary granuloma formation: analysis of cellular recruitment and cytokine responses. *Am. J. Pathol.* **154**, 1407–1416 (1999).
23. Schecter, A. D. *et al.* Tissue factor is induced by monocyte chemoattractant protein-1 in human aortic smooth muscle and THP-1 cells. *J. Biol. Chem.* **272**, 28568–28573 (1997).
24. Nibbs, R. J. B., Wylie, S. M., Pragnell, I. B. & Graham, G. J. Cloning and characterization of a novel murine beta chemokine receptor, D6. Comparison to three other related macrophage inflammatory protein-1alpha receptors, CCR-1, CCR-3, and CCR-5. *J. Biol. Chem.* **272**, 12495–12504 (1997).
25. Power, C. A. *et al.* Molecular cloning and functional expression of a novel CC chemokine receptor cDNA from a human basophilic cell line. *J. Biol. Chem.* **270**, 19495–19500 (1995).
26. Imai, T. *et al.* The T cell-directed CC chemokine TARC is a highly specific biological ligand for CC chemokine receptor 4. *J. Biol. Chem.* **272**, 15036–15042 (1997).
27. Kelder, W., McArthur, J. C., Nance-Sproson, T., McClernon, D. & Griffin, D. E. Beta-chemokines MCP-1 and RANTES are selectively increased in cerebrospinal fluid of patients with human immunodeficiency virus-associated dementia. *Ann. Neurol.* **44**, 831–835 (1998).

28. Conant, K. *et al.* Induction of monocyte chemoattractant protein-1 in HIV-1 Tat-stimulated astrocytes and elevation in AIDS dementia. *Proc. Natl. Acad. Sci. USA* **95**, 3117–3121 (1998).
29. Rollins, B. J., Morrison, E. D. & Stiles, C. D. Cloning and expression of JE, a gene inducible by platelet-derived growth factor and whose product has cytokine-like properties. *Proc. Natl. Acad. Sci. USA* **85**, 3738–3742 (1988).
30. Gunn, M. D. *et al.* A chemokine expressed in lymphoid high endothelial venules promotes the adhesion and chemotaxis of naive T lymphocytes. *Proc. Natl. Acad. Sci. USA* **95**, 258–263 (1998).

Acknowledgements

We thank A. Abbas, C. Gerard, R. Ransohoff and E. Reinherz for helpful comments and advice; P. Marks, L. Clayton and C. Daly for assistance; A. Satoskar and J. David for *Leishmania major*; and the staff of the Animal Resource Facility at Dana-Farber Cancer Institute for their humane animal care. Supported by NIH grants to B.J.R., who is also a Scholar of the Leukemia Society of America and is supported by the Novartis/Dana-Farber Drug Discovery Program.

Correspondence and requests for materials should be addressed to B.J.R. (e-mail: barrett_rollins@dfci.harvard.edu).

Betaglycan binds inhibin and can mediate functional antagonism of activin signalling

Kathy A. Lewis, Peter C. Gray, Amy L. Blount, Leigh A. MacConell, Ezra Wiater, Louise M. Bilezikjian & Wylie Vale

Clayton Foundation Laboratories for Peptide Biology,
The Salk Institute for Biological Studies, La Jolla, California 92037, USA

Activins and inhibins¹, structurally related members of the TGF- β superfamily of growth and differentiation factors², are mutually antagonistic regulators of reproductive and other functions^{1,3}. Activins bind specific type II receptor serine kinases (ActRII or IIB)^{4–6} to promote the recruitment and phosphorylation of the type I receptor serine kinase, ALK4 (refs 7–9), which then regulates gene expression by activating Smad proteins². Inhibins also bind type II activin receptors but do not recruit ALK4, providing a competitive model for the antagonism of activin by inhibin^{9–11}. Inhibins fail to antagonize activin in some tissues and cells, however, suggesting that additional components are required for inhibin action^{9,12,13}. Here we show that the type III TGF- β receptor, betaglycan^{14,15}, can function as an inhibin co-receptor with ActRII. Betaglycan binds inhibin with high affinity and enhances binding in cells co-expressing ActRII and betaglycan. Inhibin also forms crosslinked complexes with both recombinant and endogenously expressed betaglycan and ActRII. Finally, betaglycan confers inhibin sensitivity to cell lines that otherwise respond poorly to this hormone. The ability of betaglycan to facilitate inhibin antagonism of activin provides a variation on the emerging roles of proteoglycans as co-receptors modulating ligand–receptor sensitivity, selectivity and function^{16–19}.

Betaglycan binds TGF- β isoforms with high affinity and increases the functional interaction between TGF- β and its type II and type I signalling receptors^{2,20}. While screening for potential inhibin-binding proteins, we detected inhibin binding in cells expressing the complementary DNA encoding betaglycan. Figure 1a shows that membranes from HEK 293 cells transfected with betaglycan exhibited specific, high-affinity inhibin binding ($K_i = 0.6$ (0.5–0.9) nM, whereas membranes from cells transfected with empty vector had undetectable specific inhibin binding. In competition binding assays, the inhibin-binding affinity was quite low in cells expressing ActRII alone $K_i = 6.3$ (2.9–13.4) nM, consistent with previous results⁴, but increased approximately 30-fold in cells

expressing both ActRII and betaglycan $K_i = 0.2(0.1-0.3)$ nM (Fig. 1b). Moreover, we observed increased binding in cells co-expressing ActRII and betaglycan relative to the binding measured in cells expressing ActRII or betaglycan alone (Fig. 1a). Co-transfection and binding experiments with betaglycan and ActRIIB₂ (ref. 6) in HEK 293 cells yielded similar results to those seen with ActRII (data not shown). Activin binding was not detected in HEK 293 cells expressing betaglycan alone and betaglycan did not increase binding of activin to ActRII or ActRII expression (data not shown).

To determine whether inhibin can form a crosslinked complex with betaglycan, with or without ActRII, we affinity-labelled transfected HEK 293 cells with [¹²⁵I]-inhibin. Figure 2a shows that cells transfected with vector alone yielded no visible complexes (Fig. 2a, lane 1). However, we detected an inhibin-crosslinked complex of relative molecular mass ~75,000–85,000 ($M_r \sim 75K-85K$) in cells expressing ActRII (Fig. 2a, lane 2), consistent with previous observations^{4,7,10,11}. Cells transfected with betaglycan exhibited inhibin-crosslinked bands at $M_r \sim 110K$ and ~175K–250K (Fig. 2a, lane 3), similar in size to [¹²⁵I]-TGF- β 1-labelled betaglycan bands^{14,15} and consistent with previously observed inhibin-crosslinked complexes^{9,12}. Including excess unlabelled inhibin or TGF- β 1 prevented crosslinking of inhibin to betaglycan (Fig. 2a, lanes 4 and 6), whereas excess unlabelled activin had no effect (Fig. 2a, lane 5). The lack of affinity of activin for this proteoglycan supports the possibility that betaglycan binds the α -subunit of inhibin. The ability of TGF- β 1 to block inhibin crosslinking indicates that the binding site for inhibin overlaps with at least one of the TGF- β 1 binding sites on betaglycan.

Cells transfected with both betaglycan and ActRII, labelled with [¹²⁵I]-inhibin and then subjected to immunoprecipitation with an antibody directed against ActRII yielded bands corresponding to ActRII and betaglycan, indicating the presence of a stable complex containing betaglycan, ActRII and inhibin (Fig. 2a, lane 7). We also observed a large increase in the levels of inhibin-crosslinked complexes in these cells relative to cells expressing either receptor alone (Fig. 2a). A molar excess of unlabelled inhibin completely blocked [¹²⁵I]-inhibin labelling of ActRII and betaglycan in cells overexpress-

sing both receptors (Fig. 2a, lane 8); however, unlabelled activin A (Fig. 2a, lane 9) or TGF- β 1 (Fig. 2a, lane 10) only partially blocked inhibin crosslinking to ActRII and betaglycan—possibly because of the high stability of a putative inhibin/betaglycan/ActRII complex. A molar excess of unlabelled activin A has similarly been shown to competitively block only a fraction of [¹²⁵I]-inhibin labelling of endogenous ActRII and betaglycan in both KK-1 cells and K562 cells (data not shown).

We further tested whether inhibin could form crosslinked complexes with endogenous betaglycan, or betaglycan-like molecules, and ActRII in ovarian granulosa KK-1 cells²¹ and gonadotrope L β T2 cells²². After labelling KK-1 cells with [¹²⁵I]-inhibin and immunoprecipitating with antibodies against betaglycan or ActRII, we saw immunoreactive bands corresponding in size to the betaglycan core

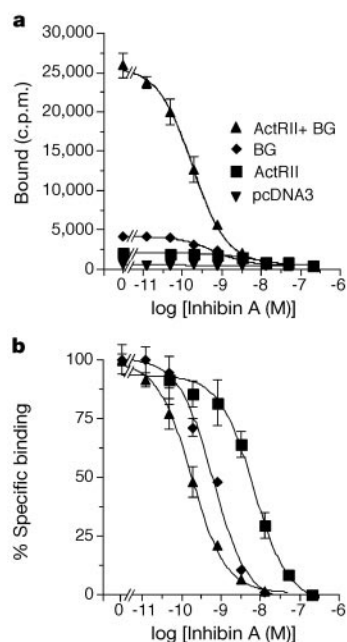


Figure 1 Betaglycan binds inhibin with high affinity and increases inhibin binding to ActRII. **a**, HEK 293 cells were transfected with ActRII-myc (ActRII), betaglycan (BG) or both (ActRII + BG), as indicated, and competition binding assays were performed. Data are presented as inhibin bound in counts per minute (c.p.m.). **b**, Data from **a**, normalized and presented as per cent specific binding. Data were analysed using the Prism software.

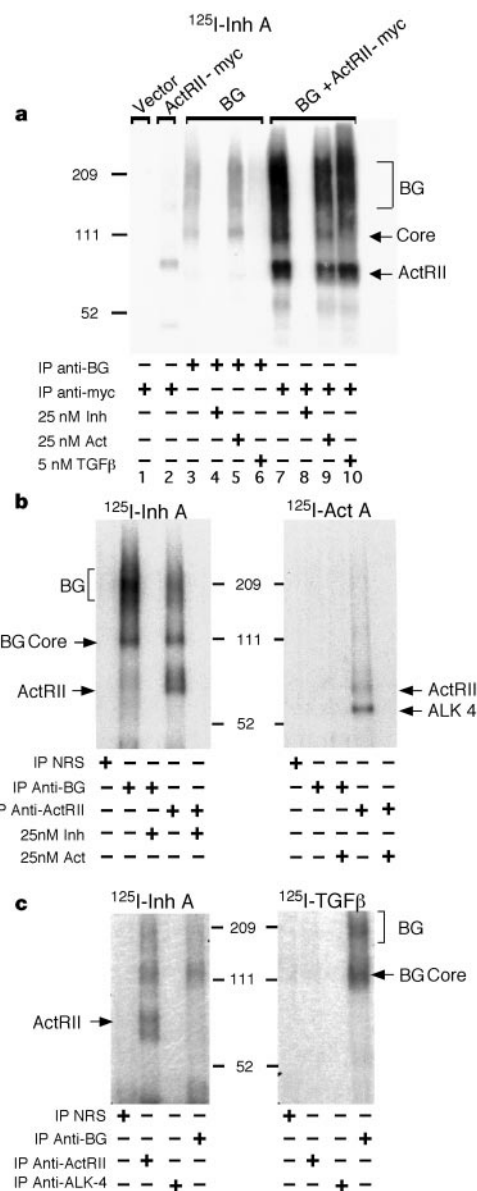


Figure 2 Inhibin forms a crosslinked complex with betaglycan and ActRII. **a**, Co-expression of ActRII and betaglycan in HEK 293 cells increases crosslinking relative to either receptor expressed alone. **b**, Inhibin A forms a complex with endogenously expressed betaglycan and ActRII in ovarian KK-1 cells. **c**, Inhibin A forms a complex with endogenously expressed betaglycan and ActRII in gonadotrope L β T2 cells. The positions of ActRII, the betaglycan core protein (BG Core), betaglycan containing glycosaminoglycan chains (BG) and ALK4 are as indicated. Molecular weight markers are represented as $M_r \times 10^{-3}$.

protein, betaglycan with glycosaminoglycan chains, and ActRII, but not the activin type I receptor ALK4 (Fig. 2b, left). Crosslinking with [125 I]-activin A showed that endogenous betaglycan does not form a covalent complex with activin in these cells (Fig. 2b, right). When extracts from [125 I]-activin-A-crosslinked KK-1 cells were subjected to immunoprecipitation with anti-ActRII antibody, bands corresponding to ActRII and ALK4, but not betaglycan, were seen (Fig. 2b, right). Gonadotrope L β T2 cells are both activin-responsive²³ and inhibin-responsive (data not shown). We labelled these cells with [125 I]-TGF- β 1 and observed crosslinked bands corresponding in size to the betaglycan core protein and glycosaminoglycan forms after immunoprecipitation with anti-betaglycan antibody but not after immunoprecipitation with normal rabbit serum (NRS) or with antibodies directed against ActRII or ALK4 (Fig. 2c, right). When crosslinking was performed with [125 I]-inhibin (Fig. 2c, left), immunoprecipitation with either anti-betaglycan antibody or anti-ActRII antibody yielded bands corresponding to betaglycan and ActRII, indicating the presence of the inhibin/betaglycan/ActRII complex. Immunoprecipitation with anti-ALK4 antibody (Fig. 2c, left) yielded no visible bands, consistent with ALK4 not being in an inhibin-labelled complex.

We explored the tissue-specific localization of betaglycan-like immunoreactivity in gonadal cell types documented to be inhibin responsive¹. In testis, we observed moderate betaglycan immunostaining in Leydig cells (Fig. 3a) and in the connective tissue of the epididymis (data not shown). The localization of betaglycan to Leydig cells is consistent with reports that inhibin is secreted by testicular Sertoli cells, acting locally to modulate steroidogenesis in Leydig cells²⁴. Positive immunostaining for betaglycan in the ovary was observed in follicular granulosa and theca cells (Fig. 3c), consistent with documented inhibin binding²⁵ and inhibin responsiveness²⁴ in these cell types, respectively. The oocyte was also strongly positive for betaglycan, in accordance with reported direct effects of inhibin on oocyte maturation²⁶ and a possible involvement of betaglycan in this process. We are currently investigating the distribution of betaglycan in the pituitary and initial results indicate that immunoreactive betaglycan is expressed in gonadotropes (L. A. M. and W. V., unpublished observations), a classic target of inhibin action.

Although many activin effects are antagonized by inhibin, there are cases in which inhibin does not block responses to activin^{9,27,28}.

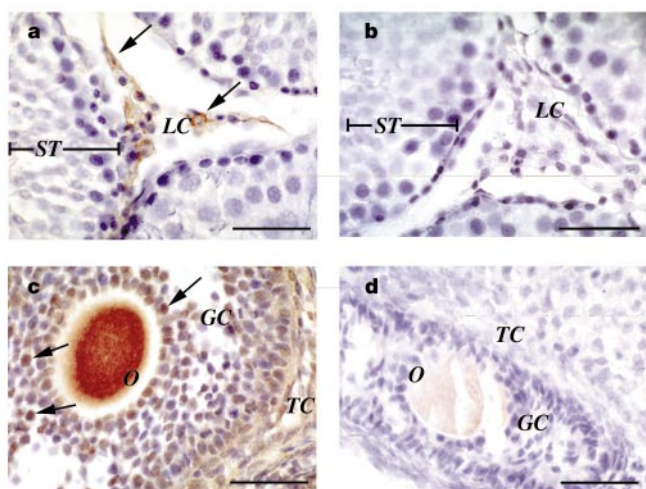


Figure 3 Immunocytochemical localization of betaglycan in normal adult rat ovary and testis. **a–d**, High magnification (60 $\times$) bright-field photomicrographs show detection of betaglycan in testis after immunostaining with anti-betaglycan antibody (**a**) or normal goat serum control (**b**), and ovary with anti-betaglycan antibody (**c**) or normal goat serum control (**d**). Representative cells immunopositive for betaglycan (brown staining) are indicated by arrows. Scale bars, 50 μ m. GC, granulosa cells; LC, Leydig cells; O, oocyte; ST, seminiferous tubule; TC, theca cells.

We previously described the ability of activin A to suppress basal ACTH secretion by AtT20 corticotrope cells and showed that inhibin has no effect on this activin response²⁸. It has been previously reported that AtT20 cells do not express detectable levels of betaglycan messenger RNA (ref. 15) and we have been unable to detect inhibin/betaglycan crosslinked complexes in this cell line (data not shown). To test directly whether betaglycan can confer inhibin responsiveness, we co-transfected AtT20 cells with betaglycan and the activin-responsive 3TPLux luciferase reporter plasmid and measured the ability of inhibin to interfere with activin signalling. In AtT20 cells transfected with empty vector, activin produced a dose-dependent increase in the resulting 3TPLux reporter activity that was unaffected by 2.5 nM inhibin (Fig. 4a). In contrast, in AtT20 cells transfected with betaglycan, this dose of inhibin substantially decreased luciferase reporter activity in response to all concentrations of activin tested (Fig. 4a). This betaglycan-dependent effect of inhibin was concentration dependent with an estimated IC_{50} of 8–10 pM inhibin (Fig. 4b). Similarly, ovarian KK-1 cells, with a weak endogenous response to inhibin (data not shown), became highly sensitive to inhibin blockade of activin signalling after transfection with betaglycan (Fig. 4c). In K562 erythroleukaemic cells overexpressing activin receptors (KAR6), activin-mediated induction of the 3TPLux reporter plasmid is largely unaffected by inhibin⁹. However, like AtT20 cells and KK-1 cells, KAR6 cells were highly sensitive to blockade of activin signalling by inhibin after transfection with betaglycan (Fig. 4d).

Several growth factors and cytokines require cell-surface proteoglycans to gain access to their respective signalling receptors and to exert biological responses¹⁶. Our data are consistent with a model (Fig. 5) in which the inhibin/betaglycan complex competes with activin for access to ActRII and therefore prevents the formation of the activin/ActRII complex required for recruitment of ALK4 and initiation of the activin-signalling cascade. This model is similar in mechanism, but not in consequence, to that proposed for TGF- β signalling in which betaglycan is thought to concentrate TGF- β at the cell surface and present it to its cognate type II receptor to

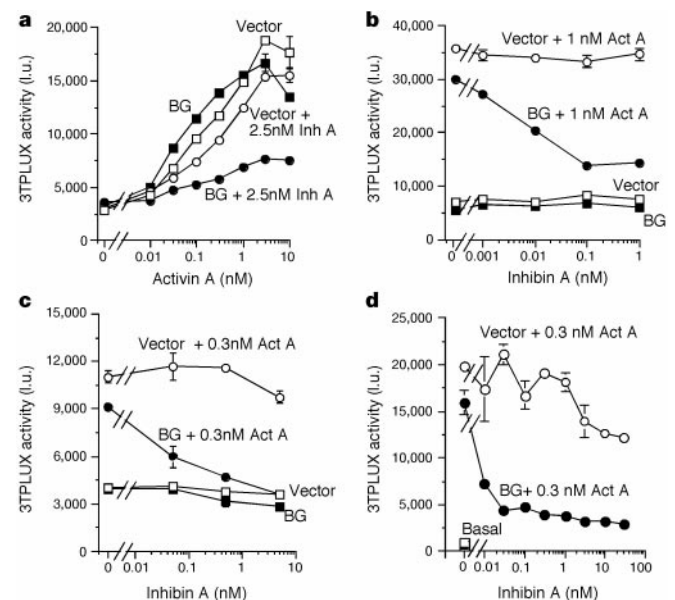


Figure 4 Betaglycan mediates functional antagonism of activin A (ActA) signalling by inhibin A (InhA) in corticotrope, ovarian and erythroleukaemic cells. **a**, Betaglycan confers inhibin responsiveness to AtT20 cells. **b–d**, The dose-dependence of the inhibin blockade of activin signalling in AtT20 cells (**b**), KK-1 cells (**c**) and IPTG-induced KAR6 cells (**d**) in the presence or absence of betaglycan (BG) is shown. Luciferase activity is represented in arbitrary luciferase units (l.u.) and normalized to β -GAL activity.

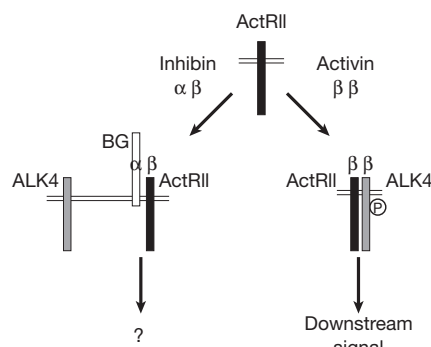


Figure 5 Diagram illustrating the proposed model in which betaglycan (BG) functions as an inhibin co-receptor with ActRII. In the presence of betaglycan, inhibin binding to ActRII is increased, access of activin to ActRII is prevented, and ALK4 is not recruited or activated.

enhance signalling²⁰. The role recently proposed for the proteoglycan *mahogany* to facilitate binding of the MSH antagonist, agouti, to the MSH receptor^{18,19} may be the most analogous to that of betaglycan on inhibin action reported here. □

Methods

Binding and crosslinking

Human embryonic kidney (HEK) 293 cells were transfected using calcium phosphate, and binding and crosslinking were performed as described²⁹. [¹²⁵I]-TGF-β1 was purchased from NEN and used in crosslinking experiments at ~1 pM per 4 × 10⁶ cells. Crosslinked complexes were isolated by immunoprecipitation using anti-betaglycan antiserum (R&D Systems, or provided by J. Massagué), monoclonal anti-myc antibody (9E10) (Calbiochem), or antibodies against ActRII or ALK4 (ref. 29).

Immunostaining

Immunohistochemistry was performed on paraffin-embedded adult rat tissues as described³⁰, using a goat anti-betaglycan serum directed against the extracellular domain of recombinant betaglycan (20 µg ml⁻¹; R&D Systems). Specificity of staining was verified using normal goat IgG or by exclusion of the primary antiserum.

Luciferase assays

Mouse A+T20, mouse KK-1 and human KAR6 cells were transfected with the 3TPLux reporter plasmid⁸, CMV-β-galactosidase (β-GAL) and either empty vector or rat betaglycan cDNA. AtT20 and KK-1 cells were transfected using Superfect Transfection Reagent (Qiagen) under optimized conditions and treated with various concentrations of activin A for 15 h in the presence or absence of inhibin A. K562-derived KAR6 cells were transfected with 3TPLux, betaglycan and β-GAL as described⁹, treated with 0.3 nM activin A in the presence of various concentrations of inhibin, and assayed for luciferase activities. All luciferase activities were normalized to β-GAL activities.

Received 21 December 1999; accepted 31 January 2000.

- Vale, W. *et al.* Chemical and biological characterization of the inhibin family of protein hormones. *Rec. Prog. Hormone Res.* **44**, 1–34 (1988).
- Massagué, J. TGF-β signal transduction. *Annu. Rev. Biochem.* **67**, 753–791 (1998).
- Matzuk, M. M. *et al.* Transgenic models to study the roles of inhibins and activins in reproduction, oncogenesis, and development. *Rec. Prog. Horm. Res.* **51**, 123–154; discussion 155–127 (1996).
- Mathews, L. S. & Vale, W. W. Expression cloning of an activin receptor, a predicted transmembrane serine kinase. *Cell* **65**, 973–982 (1991).
- Mathews, L. S., Vale, W. W. & Kintner, C. R. Cloning of a second type of activin receptor and functional characterization in *Xenopus* embryos. *Science* **255**, 1702–1705 (1992).
- Attisano, L., Wrana, J. L., Cheifetz, S. & Massagué, J. Novel activin receptors: distinct genes and alternative mRNA splicing generate a repertoire of serine/threonine kinase receptors. *Cell* **68**, 97–108 (1992).
- Attisano, L., Wrana, J. L., Montalvo, E. & Massagué, J. Activation of signalling by the activin receptor complex. *Mol. Cell. Biol.* **16**, 1066–1073 (1996).
- Carcamo, J. *et al.* Type I receptors specify growth-inhibitory and transcriptional responses to transforming growth factor beta and activin. *Mol. Cell. Biol.* **14**, 3810–3821 (1994).
- Lebrun, J. -J. & Vale, W. W. Activin and inhibin have antagonistic effects on ligand-dependent heterodimerization of the type I and type II activin receptors and human erythroid differentiation. *Mol. Cell. Biol.* **17**, 1682–1691 (1997).
- Xu, J. M., McKeehan, K., Matsuzaki, K. & McKeehan, W. L. Inhibin antagonizes inhibition of liver cell growth by activin by a dominant-negative mechanism. *J. Biol. Chem.* **270**, 6308–6313 (1995).
- Martens, J. W. *et al.* Inhibin interferes with activin signaling at the level of the activin receptor complex in Chinese hamster ovary cells. *Endocrinology* **138**, 2928–2936 (1997).
- Draper, L. B. *et al.* Identification of an inhibin receptor in gonadal tumors from inhibin α-subunit knockout mice. *J. Biol. Chem.* **273**, 398–403 (1998).

- Hertan, R., Farnworth, P. G., Fitzsimmons, K. L. & Robertson, D. M. Identification of high affinity binding sites for inhibin on ovine pituitary cells in culture. *Endocrinology* **140**, 6–12 (1999).
- Lopez-Casillas, F. *et al.* Structure and expression of the membrane proteoglycan betaglycan, a component of the TGF-β receptor system. *Cell* **67**, 785–795 (1991).
- Wang, X. F. *et al.* Expression cloning and characterization of the TGF-β type III receptor. *Cell* **67**, 797–805 (1991).
- Schlessinger, J., Lax, I. & Lemmon, M. Regulation of growth factor activation by proteoglycans: what is the role of the low affinity receptors? *Cell* **83**, 357–360 (1995).
- Tsuda, M., Kamimura, K., Nakato, H. & Selleck, S. The cell-surface proteoglycan Dally regulates Wingless signalling in *Drosophila*. *Nature* **400**, 276–280 (1999).
- Gunn, T. M. *et al.* The mouse *mahogany* locus encodes a transmembrane form of human attractin. *Nature* **398**, 152–156 (1999).
- Nagle, D. *et al.* The mahogany protein is a receptor involved in suppression of obesity. *Nature* **398**, 148–152 (1999).
- Lopez-Casillas, F., Wrana, J. L. & Massagué, J. Betaglycan presents ligand to the TGF β signaling receptor. *Cell* **73**, 1435–1444 (1993).
- Kananen, K. *et al.* Gonadal tumorigenesis in transgenic mice bearing the mouse inhibin α-subunit promoter/Simian virus t-antigen fusion gene: characterization of ovarian tumors and establishment of gonadotropin-responsive granulosa cell lines. *Mol. Endocrinol.* **9**, 616–627 (1995).
- Turgeon, J. L., Kimura, Y., Waring, D. W. & Mellon, P. L. Steroid and pulsatile gonadotropin-releasing hormone (GnRH) regulation of luteinizing hormone and GnRH receptor in a novel gonadotrope cell line. *Mol. Endocrinol.* **10**, 439–450 (1996).
- Graham, K. E., Nusser, K. D. & Low, M. J. LβT2 gonadotroph cells secrete follicle stimulating hormone (FSH) in response to activin A. *J. Endocrinol.* **162**, R1–5 (1999).
- Hsueh, A. J. W. *et al.* Heterodimers and homodimers of inhibin subunits have different paracrine action in the modulation of luteinizing hormone-stimulated androgen biosynthesis. *Proc. Natl Acad. Sci. USA* **84**, 5082–5086 (1987).
- Woodruff, T. K., Krummen, L., McCray, G. & Mather, J. P. *In situ* ligand binding of recombinant human [I-125] Activin-A and recombinant human [I-125] Inhibin-A to the adult rat ovary. *Endocrinology* **133**, 2998–3006 (1993).
- Alak, B. M., Smith, G. D., Woodruff, T. K., Stouffer, R. L. & Wolf, D. P. Enhancement of primate oocyte maturation and fertilization in vitro by inhibin A and activin A. *Fertil. Steril.* **66**, 646–653 (1996).
- Schubert, D. *et al.* Activin is a nerve cell survival molecule. *Nature* **344**, 868–870 (1990).
- Bilezikjian, L. M., Blount, A. L., Campen, C. A., Gonzalez-Manchon, C. & Vale, W. W. Activin-A inhibits proopiomelanocortin messenger RNA accumulation and adrenocorticotropin secretion of AtT20 cells. *Mol. Endocrinol.* **5**, 1389–1395 (1991).
- Gray, P. C. *et al.* Identification of a Binding Site on the Type II Activin Receptor for Activin and Inhibin. *J. Biol. Chem.* **275**, 3206–3212 (2000).
- MacConnell, L. A., Widger, A. E., Barth-Hall, S. & Roberts, V. J. Expression of activin and follistatin in the rat hypothalamus: anatomical association with gonadotropin-releasing hormone neurons and possible role of central activin in the regulation of luteinizing hormone release. *Endocrine* **9**, 233–241 (1998).

Acknowledgements

We thank K. Kunitake for assistance with binding assays and M. Perrin for helpful discussions. The rat betaglycan cDNA and anti-betaglycan antiserum were provided by J. Massagué and the KK-1 cell line was provided by A. Hsueh. The LβT2 cell line was provided by P. Mellon. This work was supported by NICHD 13527, the Kleberg Foundation, the Foundation for Medical Research, California Division and a George E. Hewitt Foundation for Medical Research fellowship to P.G.

Correspondence and requests for materials should be addressed to W. V. (e-mail: vale@salk.edu).

The Gcn5 bromodomain co-ordinates nucleosome remodelling

Popi Syntichaki[†], Irene Topalidou[†] & George Thireos^{*}

^{*} Institute of Molecular Biology and Biotechnology, FORTH, PO Box 1527, Heraklion 71110, Crete, Greece

[†] Department of Biology, University of Crete, Vasilika Vouton, 71409 Heraklio, Crete, Greece

The access of transcription factors to eukaryotic promoters often requires modification of their chromatin structure, which is accomplished by the action of two general classes of multiprotein complexes¹. One class contains histone acetyltransferases (HATs), such as Gcn5 in the SAGA complex², which acetylate nucleosomal histones. The second class contains ATPases, such as Swi2 in the Swi/Snf complex³, which provide the energy for nucleosome remodelling. In several promoters these two complexes cooperate but their functional linkage is unknown^{4–8}. A protein module that

is present in all nuclear HATs, the bromodomain, could provide such a link⁹. The recently reported *in vitro* binding of a HAT bromodomain with acetylated lysines within H3 and H4 amino-terminal peptides¹⁰ indicates that this interaction may constitute a targeting step for events that follow histone acetylation. Here we use a suitable promoter to show that bromodomain residues essential for acetyl-lysine binding are not required *in vivo* for Gcn5-mediated histone acetylation but are fundamental for the subsequent Swi2-dependent nucleosome remodelling and consequent transcriptional activation. We show that the Gcn5 bromodomain stabilizes the Swi/Snf complex on this promoter.

The yeast SAGA complex is recruited to promoters through interactions with transcriptional activators such as Gcn4 (ref. 11). The HAT activity of Gcn5 is essential for the function of SAGA (refs 12, 13) and it has been shown that Gcn5-dependent acetylation of lysine residues within the N termini of nucleosomal histones H3 and H4 is a prerequisite for transcriptional activation^{12,14}. Although much evidence supports the notion that histone acetylation is important for nucleosome remodelling, the *in vivo* consequences of this post-translational modification have not been defined because of the lack of an appropriate and simple promoter system.

To develop such a system we constructed an artificial promoter, taking advantage of the well studied nucleosomal organization and remodelling of the yeast *PHO5* promoter¹⁵. We placed a Gcn4 binding site just upstream of UAS1 (that is, within a nucleosome-free region) of the *PHO5* promoter; this hybrid promoter was used to direct the expression of the *lacZ* gene. This construct, *gPHO5*, and its predicted nucleosomal organization, is shown in Fig. 1a. The presence of the two positioned nucleosomes was verified using restriction endonuclease accessibility assays and micrococcal nuclease digestions in isolated nuclei (data not shown). The *gPHO5* promoter was expected to direct increased transcription in response to two distinct environmental signals: amino-acid starvation which induces high amounts of Gcn4 (ref. 16) and inorganic phosphate

limitation which results in the nuclear translocation of Pho4, the major activator of the *PHO5* gene¹⁵. Indeed, as shown in Fig. 1b, the expression of *gPHO5-lacZ* was low in regular minimal medium but highly induced when cells were grown either under amino-acid starvation or phosphate limitation conditions. Only the former induction was dependent on Gcn4 as well as on components of the SAGA complex such as Gcn5, Spt3 (Fig. 1b), Ada2, Spt8 and Spt20 (data not shown). The observed Gcn5 independence of Pho4 activity agreed with previous reports¹⁷. We also examined the involvement of Swi2, an essential component of the Swi/Snf nucleosome remodelling complex, and found that it is required for the Gcn4-directed activation but not for that directed by Pho4 (Fig. 1b). This analysis showed that the Gcn4-dependent expression of *gPHO5-lacZ* requires the function of both histone acetyltransferase and ATP-driven nucleosome remodelling complexes.

We next examined whether any nucleosome remodelling was observed by the recruitment of SAGA and Swi/Snf complexes. For this purpose we measured the sensitivity of DNA residing in isolated nuclei to cleavage with the restriction endonuclease *ClaI*, an assay that monitors remodelling in nucleosome -2 in the *PHO5* promoter¹⁸. As shown in Fig. 2a, sensitivity of the *gPHO5* promoter to *ClaI* cleavage was greatly increased following Gcn4- or Pho4-dependent transcriptional activation. However, the Gcn4-dependent increased sensitivity was not evident in nuclei that were derived from strains either carrying the Gcn5 F221A substitution mutant, which has no HAT activity¹², or lacking the function of the *SPT3* gene. In addition, Swi2 was required for the Gcn4-directed increase in *ClaI* sensitivity. By contrast and in agreement with previous results^{17,19}, the Pho4-dependent sensitivity of *gPHO5* DNA to this restriction enzyme was not affected by the absence of either Gcn5 or Swi2 (data not shown). We concluded that the Gcn4-specific transcriptional activation of *gPHO5-lacZ* correlates with nucleosomal remodelling that is strictly dependent on the HAT activity of Gcn5 and the Swi2 component of the Swi/Snf complex.

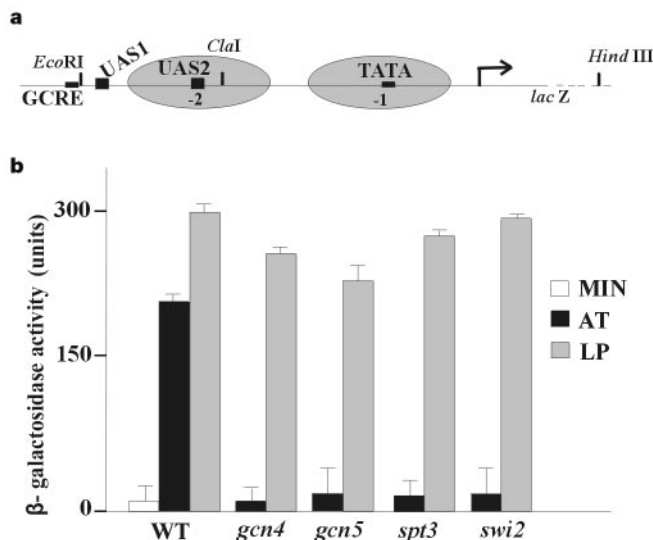


Figure 1 Expression of the *gPHO5-lacZ* reporter gene is dependent on the Gcn4 activator, the SAGA complex and the Swi/Snf nucleosome remodelling machine. **a**, Diagram of the *gPHO5* reporter indicating the positions of the transcription start site (arrow), the TATA element, the Pho4 binding sites UAS1 and UAS2 and the Gcn4-binding site GCRE. The oval circles indicate the position of nucleosomes -1 and -2. Also shown are the sites of cleavage of the restriction endonucleases we used. **b**, Measurements of β -galactosidase activity produced by *gPHO5-lacZ* reporter in the indicated strains, grown in minimal medium (MIN, white bar), and under either amino-acid starvation (AT, black bars) or phosphate limitation (LP, grey bars) conditions. Shown is the average of at least three independent experiments. WT, wild type.

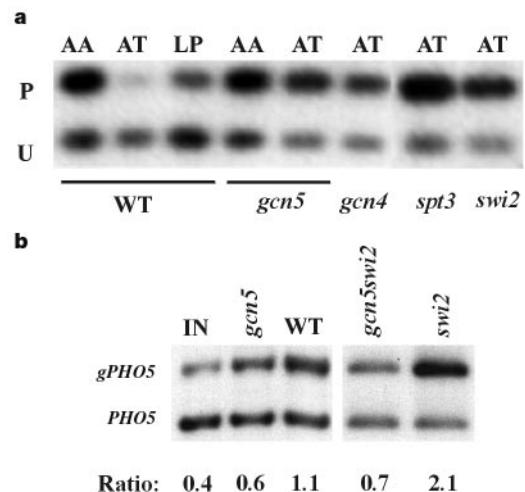


Figure 2 Remodelling of the -2 nucleosome in *gPHO5* promoter by the Swi/Snf complex requires prior histone acetylation by the SAGA complex. **a**, For determination of nucleosomal remodelling, nuclei isolated from the indicated strains grown either in rich (AA), amino-acid limiting (AT) or phosphate limiting (LP) media were digested with *ClaI*. P indicates the position of the *ClaI*-resistant *EcoRI*-*HindIII* DNA fragment (400 bp) and U indicates the *ClaI*-*HindIII*-digested DNA fragment (280 bp). The *gcn5* strain expressed the F221A allele of Gcn5. **b**, The presence of acetylated chromatin was assayed *in vivo* by measuring the ratio of the PCR-amplified *gPHO5* DNA to *PHO5* DNA (740 bp and 400 bp, respectively), present in either total (IN) or immunoprecipitated chromatin from the indicated strains. Shown is a representative of at least three independent experiments.

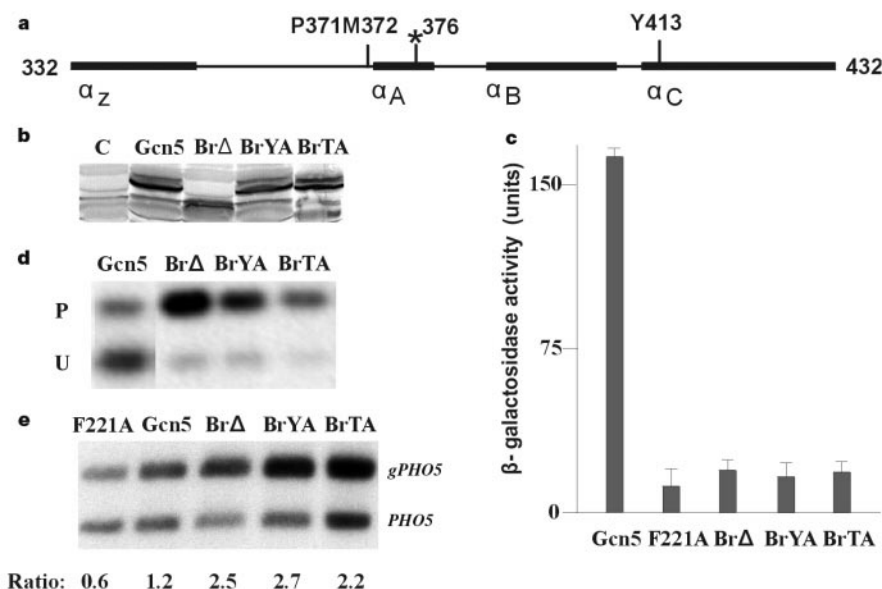


Figure 3 The Gcn5 bromodomain is required for nucleosome remodelling after histone acetylation. **a**, Diagram of the Gcn5 bromodomain indicating the four α -helices, the position of residues that were mutated and the end point (asterisk) of the deletion mutant. **b**, Immunoblot detection of haemagglutinin (HA)-tagged Gcn5 and mutant derivatives. **c**, Effects of the bromodomain mutations in the expression of *gPHO5-lacZ* in strains

containing the indicated Gcn5 derivatives and grown under amino-acid starvation conditions. **d**, Nucleosome remodelling assays in strains grown in amino-acid starvation medium and expressing the indicated Gcn5 derivatives. **e**, Chromatin acetylation assays in strains grown under amino-acid starvation conditions and expressing the indicated Gcn5 derivatives.

To determine whether Gcn5 HAT activity was independent of the nucleosome remodelling activity of Swi/Snf, we performed chromatin immunoprecipitation (CHIP) assays using antibodies raised against acetylated N-terminal peptides of histones H3 and H4. To measure hyperacetylated nucleosomes within the *gPHO5* promoter, we subjected the immunoprecipitated DNA to quantitative polymerase chain reaction (PCR) analysis using three primers that amplify both the plasmid-borne *gPHO5* DNA and the chromosomal *PHO5* region. The latter served as the most appropriate control because it corresponded to a promoter with identical nucleosomal composition to that of *gPHO5* but was repressed under the amino-acid starvation growth conditions used in our experiments. As shown in Fig. 2b, this set of primers, when used to amplify the input genomic DNA, produced a fixed ratio (0.4) of the *gPHO5* DNA to the *PHO5* DNA in all tested strains. This ratio was significantly increased (1.1) in immunoprecipitated DNA derived from a wild-type strain in a manner dependent on the HAT activity of Gcn5. This ratio was even higher (2.1) in a *swi2* strain and the increased amount of immunoprecipitated *gPHO5* DNA was also Gcn5-dependent. We concluded that, at least in this promoter, acetylation of nucleosomal histones occurs independently of the Swi/Snf complex but this modification is not sufficient for nucleosome remodelling.

As histone acetylation is a prerequisite for the remodelling function of the Swi/Snf complex on the *gPHO5* promoter, a step might exist that establishes a functional link between these activities. The recently reported *in vitro* interactions of the Gcn5 bromodomain with the N-terminal tails of histones H3 and H4 (ref. 20) and of the P/CAF with acetyl-lysine within these N-terminal peptides¹⁰ raised the possibility that such an interaction might target critical events in the pathway of nucleosome remodelling. To explore this potential we generated three mutated versions of Gcn5 (Fig. 3a): (1) deleted for part of the bromodomain (Br Δ), lacking residues 376–440; (2) an alanine substitution at Tyr 413 (BrYA); and (3) a double substitution of Pro 371 to Thr and Met 372 to Ala (BrTA). The invariant Tyr 413 in the Gcn5 bromodomain corresponds to Tyr 809 in the P/CAF bromodomain, which lies within the hydrophobic pocket that binds acetyl-lysine, and its

substitution with alanine abolishes this interaction¹⁰. Likewise, Pro 371 and Met 372 are highly conserved residues within the bromodomain family and are lined within the acetyl-lysine interacting pocket in the PCAF bromodomain¹⁰. All three mutant genes expressed stable proteins (Fig. 3b), which were unable to support transcriptional activation of the *gPHO5* reporter (Fig. 3c). Moreover, all three mutated Gcn5 proteins were unable to support nucleosomal remodelling (Fig. 3d) but CHIP analysis revealed that they supported histone acetylation on *gPHO5* nucleosomes to a level higher than that effected by the wild-type Gcn5 (Fig. 3e).

These results demonstrated that the Gcn5 bromodomain is required for nucleosome remodelling and transcriptional activation in steps subsequent to histone acetylation. Furthermore, the identical molecular and genetic phenotypes in both *swi2* and bromodomain mutant strains suggested the existence of a step requiring their coordinated function. To investigate mechanistic aspects of this coordination we monitored the recruitment of the Swi/Snf complex on the *gPHO5* promoter through CHIP analysis using an antibody against Snf5, an essential component of the Swi/Snf complex³. Figure 4 shows that Swi/Snf was associated with *gPHO5* in the wild strain. This association was dependent on Gcn4, in agreement with recent reports^{21,22}, and was evident even in the absence of Gcn5. By contrast, in the presence of the Gcn5 bromodomain mutant derivatives, Swi/Snf association was compromised. Thus, the observed compromise of nucleosome remodelling in bromodomain mutants reflected, at least, the inability of Swi/Snf to remain stably associated within the promoter.

In the simple *gPHO5* promoter, Gcn4 independently recruits both Swi/Snf and SAGA but the nucleosome-remodelling function of Swi/Snf follows the Gcn5-dependent histone acetylation. These interdependencies are not in full agreement with recent reports concerning the more complex *HO* gene promoter where it has been shown that SAGA-complex recruitment²³ and Gcn5 function¹⁴ were dependent on the previous recruitment of the Swi/Snf complex. However, as in these reports the activity of Swi/Snf was not measured, the possibility that it functions after histone acetylation cannot be excluded. Our data cannot distinguish whether the two

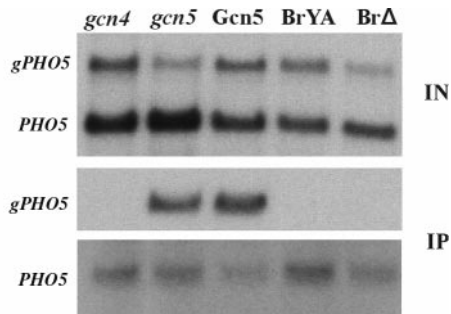


Figure 4 The Gcn5 bromodomain is required for the stable association of the Swi/Snf complex on the *gPHO5* promoter. The presence of the Swi/Snf complex on the *gPHO5* promoter was measured through PCR-amplified *gPHO5* or *PHO5* DNA present in either total (upper panel, IN) or anti-Snf5 immunoprecipitated chromatin (middle and lower panels, IP) from the indicated strains grown under amino-acid starvation conditions. *PHO5* DNA was visible following 20 times longer exposure and demonstrated the recovery of non-specific fragments in all IPs.

complexes are recruited sequentially or simultaneously on the *gPHO5* promoter, but show that the stable association of Swi/Snf with hyperacetylated nucleosomes requires functions of the bromodomain. The fact that Swi/Snf can be stably associated in the absence of Gcn5, but not in the presence of its bromodomain mutant derivatives, suggests that acetylation, although important for the function of Swi/Snf, interferes with its stabilization on the promoter. We propose that interactions of the bromodomain with acetylated H3 and H4 N termini antagonize these effects and direct nucleosome remodelling by stabilizing the Swi/Snf complex on the promoter and possibly targeting its function. Alternatively, the bromodomain might regulate the pattern of Gcn5 acetylation in a manner critical for the stable association of Swi/Snf. □

Methods

Yeast strains and growth

The WT (wild-type), *gcn5* and *ada2* strains have been described²⁴ and were isogenic to the FY strains *snf2*, *gcn5snf2*, *spt3*, *spt8* and *spt20* which were provided by F. Winston⁴. For all reported experiments, strains were grown under amino-acid limitation conditions for at least 12 h.

DNA constructs

The *gPHO5* reporter gene was constructed by inserting a PCR fragment containing *PHO5* sequences from nucleotide -392 to +3 into the *EcoRI*-*HindIII* sites of the *HIS3-lacZ* fusion gene carried on a centromeric yeast vector¹⁶. This construction placed the Gcn4-binding site into the hypersensitive region, just upstream of the positioned nucleosomes -2 and -1 in the *PHO5* promoter¹⁸. Site-directed mutagenesis of *Gcn5* was performed as described²⁴ using the mutagenic primer: 5'-CTAGCCTATTAGCATACTGTACGCGA-CGTATTCTCGCCATTG-3' for the Tyr 413 to Ala mutant (BrYA) and 5'-CCATGGTG-CTCAAGTCCGCGGTCTCTTTGATAAATCATAATAG-3' for the Pro 371 to Thr and Met 372 to Ala mutant (BrTA). The BrΔ mutant was derived as a frameshift mutation that introduced a stop codon after residue 376.

WT and Gcn5 mutants were expressed through the pYX142 yeast expression vector (Novagen) as N-terminal HA-fusions and the produced proteins were detected by immunoblot analysis as described²⁴. The allele *Gcn5F221A*, encoding a HAT-defective Gcn5 protein, was provided by D. Allis¹². β-galactosidase assays were performed as described¹⁶.

Chromatin remodelling *in vivo*

Isolation of yeast nuclei and restriction enzyme digestions were performed as described¹⁸. The extracted DNA was subjected to secondary digestions with *EcoRI* and *HindIII*, and was analysed by gel electrophoresis and DNA-blot analysis. The fragments that correspond to the plasmid *gPHO5* promoter were monitored by hybridization with radioactively labelled *EcoRI*-*HindIII* DNA fragment.

Chromatin immunoprecipitation

Chromatin was immunoprecipitated from formaldehyde fixed whole-cell extracts as described²⁵, using antiserum specific for acetylated histones H3 (at K9, K14 and K22) and H4 (at K5, K8, K12 and K16) (Upstate Biotechnology) or an antiserum against the Snf5

protein (obtained from B. Laurent). Immunoprecipitated material was eluted²⁶ and proceeded to quantitative PCR analysis, after determination of the linear range of the reaction, using the following three primers:

Primer 1: 5'-GGAATTCTGCTGACAAAGAAATA-3' which binds just upstream of the *Pho4* UAS1 in the *PHO5* promoter.

Primer 2: 5'-GGCTGCGCAACTGTTGG-3' which binds in the beginning of the *lacZ* fusion gene.

Primer 3: 5'-CATTGGCCAAAGAAGCG-3' which binds just downstream of the AUG in the coding region of *PHO5* gene.

Combination of primers 1 and 2 amplified a fragment of 740 base pairs (bp) that corresponds to the plasmid *gPHO5* promoter region, whereas combination of primers 1 and 3 amplified a 400-bp fragment that corresponds to endogenous *PHO5* promoter region. PCR reactions were spiked with ³²P-dCTP, analysed in 1.5% agarose gels and quantified by PhosphorImager.

Received 1 December 1999; accepted 4 February 2000.

- Kingston, R. E. & Narlikar, G. J. ATP-dependent remodeling and acetylation as regulators of chromatin fluidity. *Genes Dev.* **13**, 2339–2352 (1999).
- Grant, P. A. et al. Yeast Gcn5 functions in two multisubunit complexes to acetylate nucleosomal histones: characterization of an Ada complex and the SAGA (Spt/Ada) complex. *Genes Dev.* **11**, 1640–1650 (1997).
- Peterson, C. L. & Tamkun, J. W. The SWI-SNF complex: a chromatin remodeling machine? *Trends Biochem. Sci.* **20**, 143–146 (1995).
- Roberts, S. M. & Winston, F. Essential functional interactions of SAGA, a *Saccharomyces cerevisiae* complex of Spt, Ada, and Gcn5 proteins, with the Snf/Swi and Srb/mediator complexes. *Genetics* **147**, 451–465 (1997).
- Pollard, K. J. & Peterson, C. L. Role for ADA/GCN5 products in antagonizing chromatin-mediated transcriptional repression. *Mol. Cell. Biol.* **17**, 6212–6222 (1997).
- Biggar, S. R. & Crabtree, G. R. Continuous and widespread roles for the Swi-Snf complex in transcription. *EMBO J.* **18**, 2254–2264 (1999).
- Sudarsanam, P., Cao, Y., Wu, L., Laurent, B. & Winston, F. The nucleosome remodeling complex, Snf/Swi, is required for the maintenance of transcription *in vivo* and is partially redundant with the histone acetyltransferase, Gcn5. *EMBO J.* **18**, 3101–3106 (1999).
- Gregory, P. D. et al. Chromatin remodelling at the *PHO5* promoter requires SWI-SNF and SAGA at a step subsequent to activator binding. *EMBO J.* **18**, 6407–6414 (1999).
- Winston, F. & Allis, C. D. The bromodomain: a chromatin-targeting module? *Nature Struct. Biol.* **6**, 601–604 (1999).
- Dhalluin, C. et al. Structure and ligand of a histone acetyltransferase bromodomain. *Nature* **399**, 491–496 (1999).
- Utey, R. T. et al. Transcriptional activators direct histone acetyltransferase complexes to nucleosomes. *Nature* **394**, 498–502 (1998).
- Kuo, M. H., Zhou, J., Jambeck, P., Churchill, M. E. A. & Allis, C. D. Histone acetyltransferase activity of yeast Gcn5p is required for the activation of target genes *in vivo*. *Genes Dev.* **12**, 627–639 (1998).
- Wang, L., Liu, L. & Berger, S. L. Critical residues for histone acetylation by Gcn5, functioning in Ada and SAGA complexes, are also required for transcriptional function *in vivo*. *Genes Dev.* **12**, 640–653 (1998).
- Krebs, J. E., Kuo, M. H., Allis, C. D. & Peterson, C. L. Cell cycle-regulated histone acetylation required for expression of the yeast *HO* gene. *Genes Dev.* **13**, 1412–1421 (1999).
- Svaren, J. & Hörz, W. Transcription factors vs nucleosomes: regulation of the *PHO5* promoter in yeast. *Trends Biochem. Sci.* **22**, 93–97 (1997).
- Georgakopoulos, T. & Thireos, G. Two distinct yeast transcriptional activators require the function of the GCN5 protein to promote normal levels of transcription. *EMBO J.* **11**, 4145–4152 (1992).
- Gregory, P. D. et al. Absence of Gcn5 HAT activity defines a novel state in the opening of chromatin at the *PHO5* promoter in yeast. *Mol. Cell* **1**, 495–505 (1998).
- Almer, A., Rudolph, H., Hinnen, A. & Hörz, W. Removal of positioned nucleosomes from the yeast *PHO5* promoter upon *PHO5* induction releases additional upstream activating DNA elements. *EMBO J.* **5**, 2689–2696 (1986).
- Gaudreau, L., Schmid, A., Blaschke, D., Ptashne, M. & Hörz, W. RNA polymerase II holoenzyme recruitment is sufficient to remodel chromatin at the yeast *PHO5* promoter. *Cell* **89**, 55–62 (1997).
- Ornaghi, P., Ballarín, P., Lena, A. M., Gonzales, A. & Filetici, P. The bromodomain of Gcn5p interacts *in vitro* with specific residues in the N terminus of histone H4. *J. Biol. Chem.* **273**, 1–7 (1999).
- Neely, K. E. et al. Activation domain-mediated targeting of the Swi/Snf complex to promoters stimulates transcription from nucleosomal arrays. *Mol. Cell* **4**, 649–655 (1999).
- Natarajan, K. et al. Transcriptional activation by Gcn4p involves independent interactions with the SWI/SNF complex and the SRB/Mediator. *Mol. Cell* **4**, 657–664 (1999).
- Cosma, M. P., Tanaka, T. & Nasmyth, K. Ordered recruitment of transcription and chromatin remodeling factors to a cell cycle- and developmentally regulated promoter. *Cell* **97**, 299–311 (1999).
- Syntichaki, P. & Thireos, G. The Gcn5/Ada complex potentiates the histone acetyltransferase activity of Gcn5. *J. Biol. Chem.* **273**, 24414–24419 (1998).
- Strahl-Bolsinger, S., Hecht, A., Luo, K. & Grunstein, M. SIR2 and SIR4 interactions differ in core and extended telomeric heterochromatin in yeast. *Genes Dev.* **11**, 83–93 (1997).
- Kuras, L. & Struhl, K. Binding of TBP to promoters *in vivo* is stimulated by activators and requires Pol II holoenzyme. *Nature* **399**, 609–613 (1999).

Acknowledgements

We thank D. Thanos for histone antibodies, B. Laurent for the Snf5 antibody, F. Winston for strains, D. Allis for the Gcn5F221A mutant and E. Remboutsika for suggesting the BrTA mutant. This work was supported by a European Union Training and Mobility for Researchers grant.

Correspondence and requests for materials should be addressed to G. T. (e-mail: Thireos@imbb.forth.gr).

We have the biotechnology

From insects to FISH, more from companies with designs on your lab space.

Roller-Blot HB-3D

Technique www.technique.co.uk

Compact blotting, around the compass

Most blotting techniques, including Southern and western using either radioactive or fluorescent probes, can be accommodated in the Roller-Blot HB-3D Hybridisation Incubator. This bench-standing unit accommodates a full range of hybridization vessels in an easy-viewing heated chamber. Each unit holds either three of the large 80 mm diameter × 240 mm hybridization tubes, nine of the small 44 × 220 mm tubes, or 12 mini 32 × 205 mm tubes. Alternatively, a combination of the three sizes can also be accommodated.

Reader Service No. 100



Roll playing: Roller-Blot keeps tubes in view.

CCD cameras

UVItec Limited www.uvitec.demon.co.uk

Economical imaging

These two cameras, UVIsnap and UVIsave, are designed to simplify viewing and printing of gel images. UVIsnap is a monochrome CDD camera with zoom and close-up lens and UV/IR interference filter. Mounted on a thermo-formed hood that can be placed on a UV or white light transilluminator, UVIsnap comes with a 4-inch LCD display, camera integration with 10 pre-set values, live and freeze controls and a thermal printer port, all located on the camera control box. UVIsave incorporates all the features of UVIsnap with the addition of a 3.5-inch floppy disk drive and PC software for saving images as TIFF files. The software also includes image manipulation, molecular-weight calculation, quantification and colony-counting functions.

Reader Service No. 101



Good save: TIFFs on-disk from UVIsave.

LightCycler

Roche Molecular Biochemicals

<http://biochem.roche.com/lightcycler>

Quantitative, flexible PCR

Recent improvements to the LightCycler instrument and software allow simultaneous detection of multiple targets in multiplex PCR using fluorescein, LightCycler Red 640 and LightCycler Red 705. It takes just 30 minutes to perform 40 PCR cycles, and complex quantification and mutation analysis experiments can be carried out online in real time. Cross-talk compensation eliminates signals from neighbouring channels.

Reader Service No. 102

Vectastain ABC-AmP

Vector Laboratories www.vectorlabs.com

The ABC of western blots

The Vectastain ABC-AmP Western Blot Detection System, say manufacturers Vector Laboratories, offers better sensitivity and lower background for blotting than previously available western blot detection kits. Based on Vectastain ABC (avidin-biotin complex) technology, the system uses an amplified ABC-alkaline phosphatase complex optimized for use on nitrocellulose or PVDF membranes. When used with the DuoLux chemiluminescent/fluorescent substrate, the system produces a high signal-to-noise ratio and sustained light emission. Film can be exposed to blots many hours after substrate development, making it possible to optimize band intensities or resolution. ABC-AmP kits are also available with calorimetric substrate BCIP/NBT.

Reader Service No. 103

Streptavidin-coated plates

Nalge Nunc International

www.nalgenunc.com

Convenient formats for a frequently used assay technique

Nunc offers a streptavidin-coated surface in a solid, transparent, 96-well plate format or as eight well strips, in C96 and C8 format. Streptavidin's high affinity for biotin and biotinylated molecules makes a streptavidin-coated surface ideal for binding small molecules such as peptides, oligonucleotides, haptens and other biotinylated biomolecules. Streptavidin-coated plates are used in DNA hybridization assays, immune assays and other solid-phase assays.

Reader Service No. 104

TransIT-Insecta

PanVera

www.panvera.com

High-efficiency transfection in insect cell lines

The TransIT-Insecta Transfection Reagent, developed by PanVera subsidiary Mirus Corporation, is a lipid mixture specifically designed for superior transfection efficiencies in insect cell lines. The reagent is effective in expressing genes in insect cells and transfecting Sf9, Sf21 or High Five cells with recombinant baculovirus expression systems.

Reader Service No. 105

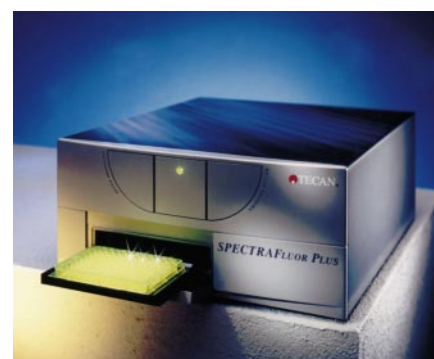
Spectrafluor Plus

Tecan

www.tecan.com

Multifunctional microplate reader

Tecan's Spectrafluor Plus is a multifunctional microplate designed to replace several instruments. It also integrates into robotic systems for full automation. Suggested applications include cell and metabolic assays and protein or DNA quantification, using any plate format up to 1,536 wells. Time to read a 96-well plate is under 30 seconds. The Spectrafluor Plus measures fluorescence or absorbance from



Multiple choice: Spectrafluor Plus likes robots.



Dedication: GeneQuest comes pre-programmed.

UV to near-IR wavelengths, as well as glow luminescence, and can be quickly switched between the different modes of measurement.

Reader Service No. 106

Triathler

Hidex

www.hidex.com

A portable multilabel tester

The Triathler is a small, manual, single-well instrument that gives instant results in liquid scintillation counting, gamma counting and luminescence counting. Due to its small size (weighing 9 kg) and portability, the Triathler is especially useful in applications that require on-the-spot measurements, such as contamination control and environmental monitoring. Bioassays and tracer measurements can be performed using biodegradable cocktails available from Hidex. Software is available to edit and save results, which can then be transferred to Excel.

Reader Service No. 107

GeneQuest Analyser

Cecil Instruments www.cecil-instruments.co.uk

A spectrophotometer for molecular biology

Fully programmed methods in the GeneQuest Analyser include those for RNA, dsDNA, ssDNA, oligonucleotide primer, proteins, bacterial growth and cell culture measurements and molecular calculations, including

theoretical thermal melt. Programmed methods are selected from a menu and displayed with headers, prompts and details on the large display screen. All wavelengths for any given assay are automatically selected and measured. Results are displayed on the screen and a print-out, timed and dated, is available on an external or an internal printer.

Reader Service No. 108

MAX-BAX

Charles River Therion www.theriondna.com

A new DNA-based testing service

The MAX-BAX DNA-profile test for marker-assisted accelerated backcrossing, can help speed up the development and production of animal research models. Results from this test can be used for the directed selection of individuals/breeders with preferred genetic backgrounds. By breeding selectively rather than randomly, an investigator can significantly decrease the number of backcross generations needed to produce a 95–99% congenic strain.

Reader Service No. 109

Lyse-N-Go

Pierce Chemical

www.piercenet.com

Lyse and amplify in a single tube

The ready-to-use Lyse-N-Go PCR reagent provides PCR-ready DNA from cells or tissue in 10 minutes or less. The reagent is recommended for high-throughput screening and other applications, as it allows the direct use of lysed samples for amplifying. Lyse-N-Go is available in packs of five 2-ml vials.

Reader Service No. 110

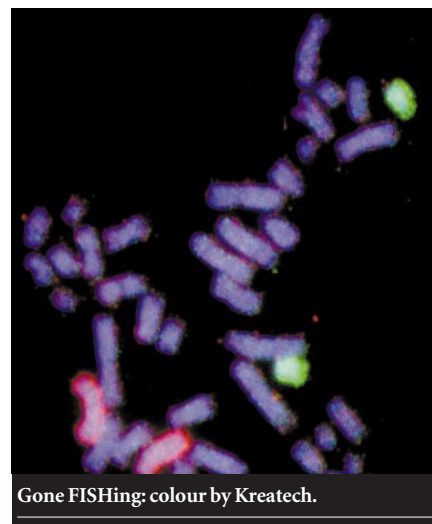
ULYSIS labelling systems

Kreatech

www.kreatech.com

Rhodamine and dGreen labelling

The ULYSIS family of nucleic-acid labelling systems now includes direct fluorescence nucleic-acid labelling kits for rhodamine and dGreen, joining kits for biotin, fluorescein and dinitrophenol, already in the range. The kits are based on Kreatech's ULS labelling system.



Gone FISHing: colour by Kreatech.

ULS-labelled probes have numerous applications in spot blot hybridization, RNA or DNA *in situ* hybridization and also (multicolour) fluorescent *in situ* hybridization (FISH).

Reader Service No. 111

FailSafe PCR

Epicentre

www.epicentre.com

Overcoming difficulty

The FailSafe PCR System has been developed to provide high-fidelity DNA amplification for routine and 'difficult' templates. Both long (at least 20 kb) and difficult templates (such as sequences with high G+C or a high degree of secondary structure) can be amplified consistently using this system. The System includes twelve PreMixes that contain dNTPs, buffers and various concentrations of MgCl₂ and FailSafe PCR Enhancer (with betaine). The user simply adds template, primers, and the FailSafe PCR Enzyme Mix to the PreMixes and then amplifies to determine the optimal PreMix.

Reader Service No. 112

These notes are compiled in the Nature office from information provided by the manufacturers. For more details, fill in the reader service card.

BLOCKERS

K⁺ Channel Blocker

E4031

Arachidonic Acid Cascade Blocker

AA 861

Wako USA

877-714-1921

Wako

Wako GmbH

(49) 02131 3110

www.wako-chem.co.jp

READER ENQUIRY NO. 35

Check out our homepage at <http://www.resgen.com>

Peptides

\$14/residue
10 residue minimum

Polyclonal Antibodies - \$1,195

Please contact us about complete services for your custom peptide and polyclonal antibody needs.

Research Genetics, Inc.
U.S. or Canada 800-533-4363
U.K. 0-800-89-1393
FAX 256-536-9016

READER ENQUIRY NO. 4

Database Management Systems

Bioinformatics

Powered by GCG

► Comprehensive analysis packages

- Enterprise-wide solutions
- Extendable software
- Computing security
- Scriptable analysis

www.gcg.com

GCG is an Oxford Molecular Company

Client/Server & Desktop Software

READER ENQUIRY NO. 36